Vol. 39, No. 4 August, 1949

Journal of Experimental Psychology

EDITED BY

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CLARENCE H. GRAHAM, *Columbia University*
HARRY HELSON, *Bryn Mawr College*
DONALD B. LINDSLEY, *Northwestern University*
DONALD G. MARQUIS, *University of Michigan*
GEORGE A. MILLER, *Harvard University*
L. L. THURSTONE, *University of Chicago*
G. R. WENDT, *University of Rochester*
JOSEPH ZUBIN, *N. Y. Psychiatric Institute, Columbia University*

CONTENTS

Electrical Responses of the Human Retina: L. A. RIGGS AND E. P. JOHNSON 415
The Effect of Inter-Sensory Stimulation on Dark Adaptation and Night Vision: A. CHAPANIS, R. O. ROUSE, AND S. SCHACHTER 425
A Scale of Subjective Brightness: R. M. HANES 438
Visual Resolution with Periodically Interrupted Light: VIRGINIA L. SENDERS 453
Response Latency as a Function of the Amount of Reinforcement: D. ZEAMAN 466
Factors Influencing the Occurrence of Reminiscence: Attempted Formal Rehearsal during the Interpolated Period: J. H. ROHRER 484
Reminiscence in Pursuit-Rotor Learning as a Function of Rest and of Amount of Pre-Rest Practice: A. L. IRION 492
Performance and Reminiscence in Motor Learning as a Function of the Degree of Distribution of Practice: G. A. KIMBLE 500
Extinction as a Function of Partial Reinforcement and Distribution of Practice: VIRGINIA F. SHEFFIELD 511
Distributed Practice in Verbal Learning and the Maturation Hypothesis: SUSAN T. H. WRIGHT AND D. W. TAYLOR 527
Ability Patterns under Distributed Practice: MARY J. KIENTZLE 532
The Perception of the Vertical: I. Visual and Non-Labyrinthine Cues: C. W. MANN, N. H. BERTHELOT-BERRY, AND H. J. DAUTERIVE, JR. 538
The Apparent Length of Tilted Lines: W. C. SHIPLEY, BARBARA M. NANN, AND MARY JANE PENFIELD 548
The Relative Difficulty of the Number, Form, and Color Concepts of a Weigl-Type Problem: D. A. GRANT, O. R. JONES, AND BILLIE TALLANTIS 552
Involuntary Blink Rate and Illumination Intensity in Visual Work: M. A. TINKER 558
The Effect of a Fixated Figure on Autokinetic Movement: R. S. CRUTCHFIELD AND W. EDWARDS 561
The Discrimination of Small Differences in the Time of Mechanical Stimulation: F. L. SMITH, A. L. SWEET, AND N. R. BARTLETT 569
'Spread of Effect' without Reward or Learning: F. D. SHEFFIELD 575

PUBLISHED BI-MONTHLY BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
PRINCE AND LEMON STS., LANCASTER, PA.
AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879
Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. & R., authorized September 11, 1947.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $7.00 (foreign, $7.25), single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Ave., N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Articles of a length exceeding 25 printed pages cannot be accepted.

With rare exceptions, articles are published in the order of their receipt. Authors are supplied with 50 free offprints, without covers, in the case of articles of seven or more printed pages, and with five free copies of the issue in the case of shorter contributions. One-half of the cost of tables, cuts, formulas, and author's alterations in galley proof is charged to the author.

Immediate publication can be given to articles whose authors assume the full cost of publication.

Address articles submitted for publication to the editor, Francis W. Irwin, Department of Psychology, University of Pennsylvania, Philadelphia 4, Pa.

This JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. ANDERSON and W. L. VALENTINE, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *Es*, *E*'s, and *Es*'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts should be typed double-spaced throughout, *including tables, footnotes, and bibliography.* Tables, figures, and footnotes should be placed on separate pages at the end of the manuscript.

Journal of
Experimental Psychology

VOL. 39, No. 4 AUGUST, 1949

ELECTRICAL RESPONSES OF THE HUMAN RETINA [1,2]

BY LORRIN A. RIGGS

Brown University [3]

AND

E. PARKER JOHNSON

Bowdoin College

INTRODUCTION

For over 80 years it has been known that a momentary change in electrical potential arises in the retina of the eye upon stimulation by light. The usual method of recording this difference of potential has been to make contact with the anterior and posterior poles of the eyeball. Early work on animal subjects, ranging from arthropods to man, has established many facts in connection with this electrical response, but has failed to reveal its exact nature or site of origin.

Experiments on electrical responses of the human retina are nearly as old as those on animals. An obvious difficulty with using human subjects is the danger of painful irritation while making an electrical circuit involving the cornea and some neutral point on the head. In 1941 (5) a new method was developed for recording electrical responses from the human retina. A contact lens was used to provide a stable electrical connection with the cornea, whose level of electrical potential is known to be governed principally by the potential difference across the retina. It is possible by this means to make continuous and reproducible records of the retinal action potentials developed in response to stimulation of the eye by light.

Johnson (3) has used the contact-lens method to make a quantitative study of the effects of dark adaptation. He has expressed his data in terms of curves relating magnitude of response to time in the dark. These curves continue to rise over a period of an hour or more, accompanying the rise in visual sensitivity as dark adaptation proceeds.

Adrian (1, 2) has employed the usual moist-thread electrode in a comprehensive series of qualitative experiments on the retinal responses of cats,

[1] A first report of the experiments reported in this paper was read at a meeting of the National Academy of Sciences in Washington, D. C., November 19, 1947.

[2] The experiments were supported by a grant to one of the authors (L. A. R.) from the Penrose Fund of the American Philosophical Society.

[3] The experiments reported in this paper were performed in the Psychological Laboratories of Brown University in 1945–47.

rabbits, guinea pigs, pigeons, monkeys and men. Adrian has established the fact that for the eyes of pigeons, monkeys and men, there are two principal components of the retinal response wave. The first component is rapid and diphasic. It is more apparent in response to light of the longer wavelengths and in the light-adapted eye. The second component is slower and monophasic. It is more readily aroused by light of the shorter wavelengths and in the dark-adapted eye. The first component has been attributed to photopic mechanisms, and is absent from the records of animals having few cone receptors (cat, rabbit) or none at all (guinea pig). The second, or scotopic, component appears in responses from all of the animals tested, though its magnitude and latency may be reduced by light adaptation to such an extent that it resembles the first component.

The purpose of the present experiments is to provide a series of quantitative determinations of the scotopic component of the human retinal response. To that end, a form of recording has been used which gives an undistorted record of the response wave. The fully dark-adapted eye has been used to obtain maximal responses. The reduced responses during light adaptation have also been recorded. The specific sensitivity of the eye to lights of various wavelengths has been measured in order to compare this function with the scotopic luminosity curve as obtained by the usual psychophysical procedures. The results obtained in all of these experiments support the conclusion that the major component of the human retinal response is scotopic in its characteristics. Minor differences in wave form have been noted, however, in responses to lights of various wavelengths. These differences have been attributed by Adrian (1) to differences in the relative contributions of the scotopic and photopic systems.

Apparatus

The apparatus used in these experiments has been described in detail by Johnson (3). Briefly, it consists of an optical system for stimulating the eye, recording electrodes, an amplifier, and a photographic recording device.

Optical system

Light adaptation is secured by the use of a hemispheric field that may be completely illuminated by white light of any desired intensity. This adapting surface covers most of the visual field. The test stimuli appear as added flashes of light at the center of the adapting field. Each flash has a duration of 1/25 sec. and is seen as a circular patch 7°30′ in diameter. A red fixation-point assures central stimulation. By the use of filters it has been possible to vary the intensity and wavelength characteristics of the stimulus flash. Farrand interference-type filters and Wratten 'monochromatic' filters have provided stimuli which are confined to relatively narrow bands of the spectrum at selected points throughout the visible range from 400 to 700 mμ.

The radiant energy delivered to the eye by any given flash of light is represented by the integral of the energies at all wavelengths throughout the visible spectrum. This has been computed, for each of the flashes used in these experiments, by finding the area under a curve relating energy level to wavelength. The procedure used in computing this energy will be described more fully in a forthcoming article (6).

Recording electrodes

A plastic contact lens was fitted by an optometrist to the right eye of the subject. A hole was drilled through the corneal portion of the contact lens at a point slightly nasal to the center. The hole was plugged with a disc of silver, to whose outer surface a flexible wire was attached in the manner described by Riggs (5). The inner surface was coated with a layer of silver chloride for use as the critical electrode. This surface was in direct contact with the physiologically isotonic fluid which was used to fill the space between the contact lens and the cornea. Hence, a constant, diffuse electrical connection was maintained with the anterior portion of the eye. A neutral electrode (which is also a chlorided silver disc) made similar connection with a fluid in contact with the forehead at a point above the right eye. It is the differ-

ence of potential existing between these two electrodes which is amplified and recorded.

Amplifier and recording device

Leads from the recording electrodes are taken to a high-gain, direct-coupled amplifier. The output of this amplifier is used to actuate the first element of the recording device, a General Electric loop oscillograph. A beam from the mirror of this oscillograph makes a trace upon a moving strip of photographic paper. The resulting records show the exact magnitude and wave-form of the retinal action potentials built up in response to the flashes of light.

An important feature of the apparatus is its ability to make a faithful record of even quite slow potential changes. Adrian has pointed out that the scotopic component is of such long latency and duration that it is somewhat distorted by the type of resistance-capacity coupling that he was obliged to employ. The use of a contact lens in the present experiments has resulted in sufficient stability and freedom from electrode drift so that direct coupling has always been employed.

Procedure

The general procedure observed in all experimental sessions was as follows: The electrodes were applied, and appropriate connections were made to the amplifier. After a period of dark adaptation, the white adapting field was turned on and adjusted to the desired level of brightness. (This step was, of course, omitted in the experiments on the dark-adapted eye. In the latter experiments the period of dark adaptation preceding the test flashes was 40 min.) In the experiments on levels of light adaptation, a schedule was followed which provided at least 15 min. in the dark followed by adaptation to the appropriate level of brightness of surround. A minimum of 15 min. of exposure to this surround was provided before the series of stimulus flashes was begun.

Preliminary tests indicated that it was permissible to deliver the brief flashes employed in this experiment at a rate of approximately one flash per min. for the weaker flashes and one every two min. for more intense ones. At this rate there was no danger of a decline in the sensitivity of the eye. The duration of each experimental session was about three hours. It was necessary to combine the results of several experimental sessions to obtain the desired coverage of intensities and wavelengths of stimulation for each level of adaptation. A very gratifying agreement was found from day to day in responses to a given stimulus flash. Two young men with normal acuity and color vision served as subjects in these experiments.

Results

Responses to white light

Wave-forms of responses at various levels of light-adaptation.—Fig. 1 shows a series of responses to flashes of white light, as traced from portions of our experimental records. The five responses at the extreme right of the figure were obtained after 50 min. of dark adaptation. The uppermost of these five responses was elicited by a flash having a brightness equivalent to approximately one and one-half million footlamberts.[4] This flash produced a very large *a*-wave (downward deflection of short latency) followed by a *b*-wave (first upward deflection). The next lower record, elicited by a flash having one-tenth as much energy, exhibits a smaller *a*-wave. At still lower energies of flash, the *a*-wave disappears and the *b*-wave declines. The *b*-wave meanwhile is characterized by an increasing latency and decreasing magnitude until finally no response occurs to flashes much less intense than those represented in this figure.

The remainder of the responses in Fig. 1 were obtained while the subject's retina was being flooded with light from the hemispheric adapting field. It is to be noted that the magnitude of the response is markedly reduced as the adaptation level is raised. Curiously enough, this decrease in magnitude of response is now accompanied by a decrease in its latency. With a reduction in the intensity of the stimulating flash at any one level of adaptation the more

[4] This brightness was computed by a method described by Johnson (3). It represents the photometric brightness of a field, viewed with the natural pupil, which would appear to be of the same brightness as this stimulus flash.

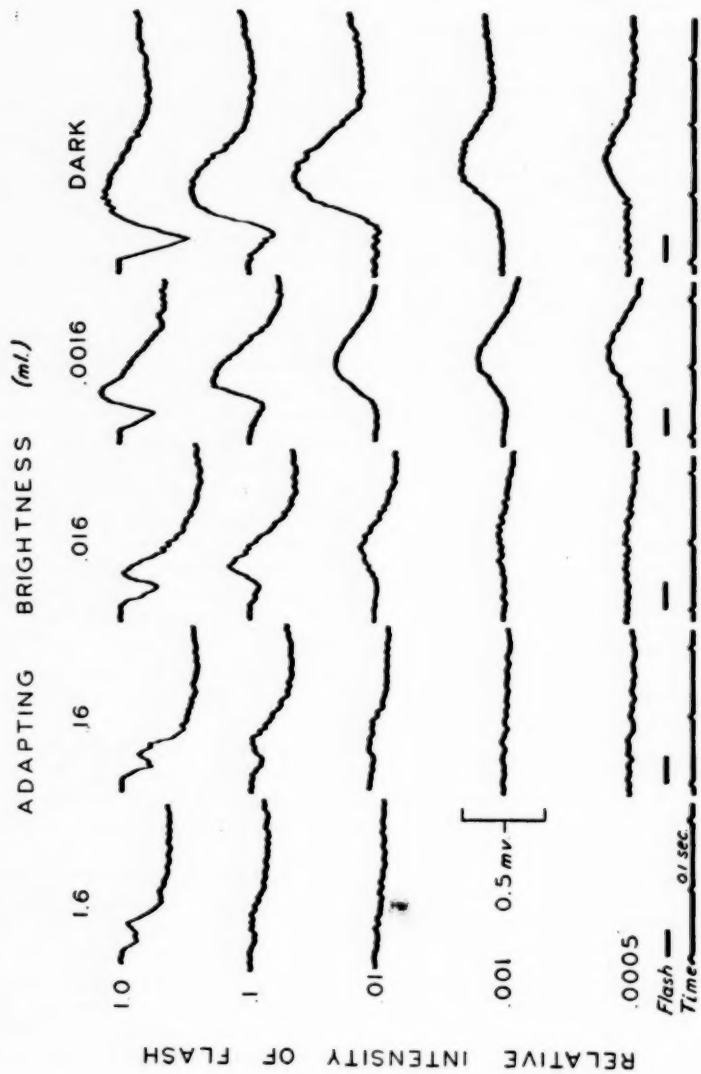

FIG. 1. Tracings of oscillographic records of human retinal responses at each of five levels of adaptation. Each record shows the initial (a- and b-wave) portion of the response to a flash having a duration of 40 msec. The highest (1.0) intensity of flash corresponds to a brightness of approximately 1.5×10^8 footlamberts.

usual relationship (decreasing magnitude of response with increasing latency) obtains. Another point which may be noted in connection with Fig. 1 is that there is a rough correspondence between the level of adapting brightness and the energy of a flash which is required to elicit a given height of response.

Perhaps the most surprising result shown by Fig. 1 is that even moderate levels of light adaptation have almost obliterated the responses to very intense flashes of light. At a level of 1.6 millilamberts (ca. 1.5 footlamberts), for example, a flash almost 10,000 times as bright (the one labelled .01 in the figure) produces no measurable response.

Magnitudes of responses at each adapting level.—Fig. 2 presents some measurements of *b*-wave magnitudes for responses similar to those of Fig. 1. These measurements, particularly for the larger magnitudes of response, are somewhat arbitrary in the sense that the height of the *b*-wave is obviously cut down by the *a*-wave which precedes it. Some attempt has been made to allow for this factor by measuring from the trough of the *a*-wave to the crest of the *b*-wave. It is probable, however, that there is no simple significance to the upper portions of these curves. Nevertheless, the curves of this figure serve to show again the large extent to which even moderate amounts of light adaptation have cut down the magnitude of the response.

Reduction in sensitivity of the eye.—It is obvious from Figs. 1 and 2 that the sensitivity of the eye has been progressively reduced by successively higher levels of light adaptation. We may derive a curve for sensitivity as a function of adapting level in the following manner: A response magnitude of 58 microvolts is arbitrarily chosen as an effect which may be produced by an appropriate intensity of flash at each level of light adapta-

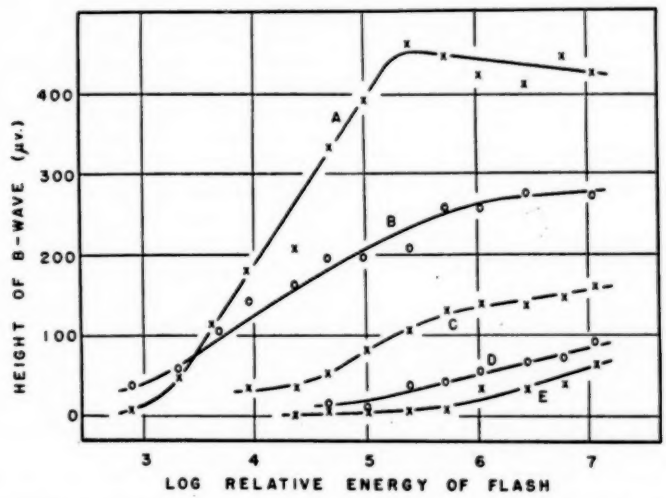

FIG. 2. Magnitudes of responses at each of five levels of adaptation. Curve A is for the dark adapted eye. Curves B, C, D and E are for adaptation levels of .0016, .016, .16, and 1.6 millilamberts respectively.

tion. (Some of the considerations involved in the choice of this level of response have been described by Johnson, 3.) It may be noted that responses of this magnitude are sufficiently large to be clearly measurable on the records, yet small enough so that they consist principally of the *b*-wave, the 'scotopic' component.) At high levels of adaptation, a relatively high intensity of flash is required to produce this effect, and the eye may be said to have reached a relatively insensitive condition. At low levels, and particularly for a state of complete dark adaptation, the reverse is true.

Following the above line of reasoning, we may define the 'sensitivity' of the eye, measured electrically, as the reciprocal of the energy of flash necessary to arouse a response of 58 microvolts. Fig. 3, derived from the curves of Fig. 2, shows a log-log plot of this 'sensitivity' against level of light adaptation.[5]

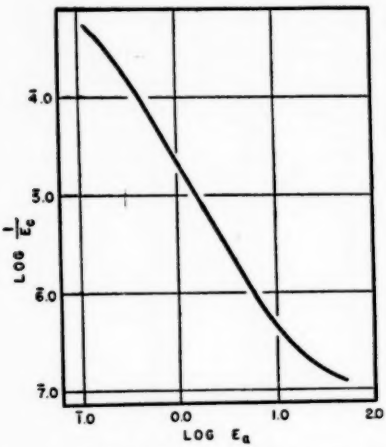

FIG. 3. Retinal 'sensitivity' as a function of level of light adaptation. Ordinate values represent log $1/E_c$, where E_c is the relative energy of flash necessary to arouse a 58-microvolt response at each adaptive level as indicated in Fig. 2. Abscissa values are log relative energy of adapting light, E_a reaching the retina. These values are computed from the brightness levels of Fig. 2 by correcting for diameter of pupil at each level of brightness. Thus, the adapting brightness of 1.5 foot-lamberts corresponds to log $E_a = 1.71$.

Responses to colored lights

Wave-forms specific to the various colors.—Responses to restricted spectral bands of light have characteristic wave-forms. Fig. 4 shows typical responses of the dark-adapted eye to red, green, and blue stimuli. While the differences are slight, they are consistent from experiment to experiment and from one subject to another. It is seen that the red light arouses an initial quick dip and rise (diphasic, 'photopic' component of Adrian) which is not present in the other records. Responses to blue light are consistently free from this sort of initial effect, while responses to intermediate colors (yellow, green) show slight traces of it from time to time.

Light adaptation produced changes in wave-form which were similar to those for white light. It was not possible in these experiments, however, to secure sufficiently intense lights of all colors to arouse responses at the higher levels of light-adaptation.

Magnitudes of response to each color.—Fig. 5 shows the magnitudes of the principal (*b*-wave) component of the response to lights of various dominant wave-lengths. This figure is for the dark-adapted eye. It is seen that the shapes of some of these curves are not unlike that for white light (Curve A in Fig. 2). It is obvious, however, that blue and green lights are highly

[5] For the purpose of Fig. 3, it has been necessary to convert the units of adapting brightness into the same units of relative energy which were used to specify the intensity of the test flash.

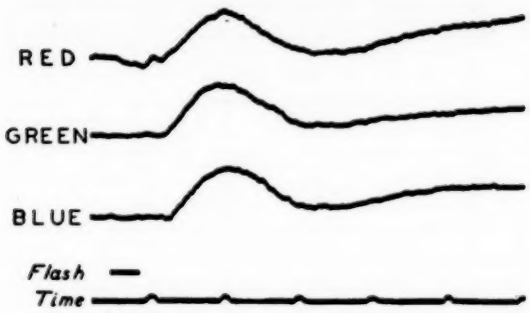

FIG. 4. Responses of the dark-adapted eye to red, green, and blue flashes of light

effective in eliciting the response, while yellow, orange, and red lights are progressively less effective.

The effectiveness of each color in arousing a given level of response (58 microvolts, as before) is shown in Curve A of Fig. 6. Values on the ordinate of this curve are derived from Fig. 5 and represent relative energies, in logarithmic units, to evoke the 58-microvolt response. Abscissa values are the wavelengths at which most of the energy is concentrated for a given filter combination.

Curve B in Fig. 6 is a scotopic luminosity curve obtained by a conventional psychophysical procedure. The procedure was as follows: The filtered light emerging from the optical system was allowed this time to diverge from the concentrated point normally occupied by the pupil of the

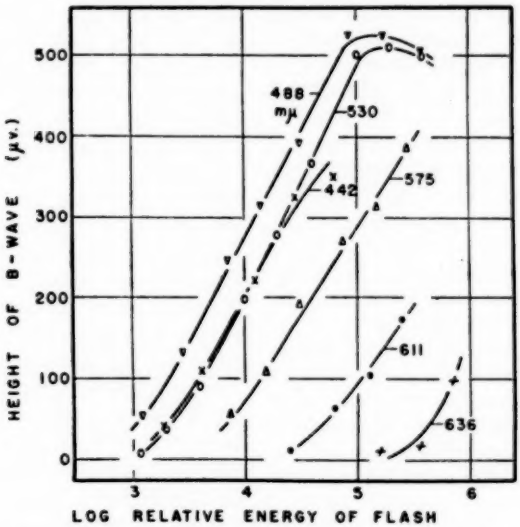

FIG. 5. Magnitudes of responses of the dark-adapted eye to lights of various dominant wavelengths

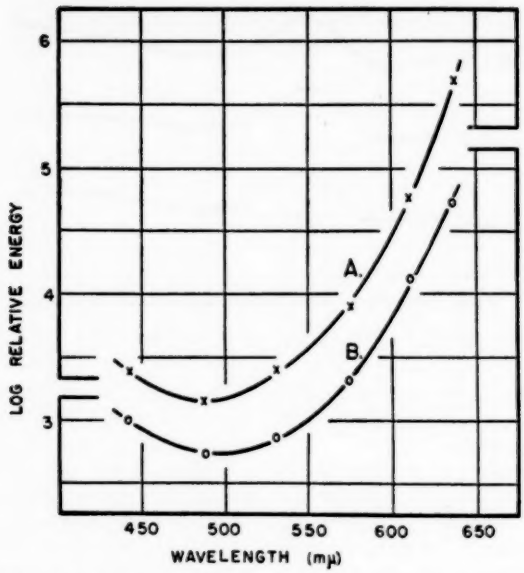

FIG. 6. Comparison of electrical and psychophysical data on the relative effectiveness of lights of various wavelengths. Curve A shows the relative energies of light required to elicit an electrical response of 58 microvolts. Curve B shows the relative energies of light required to obtain a brightness of approximately 10 microfootlamberts in the dark adapted eye.

eye, and to fall on a white screen (MgO surface) several feet away. The brightness of the screen so illuminated was in the neighborhood of 10 microfootlamberts, or not far above the absolute threshold for the dark-adapted eye. A semi-circular patch of this light was roughly equated in brightness to an adjacent standard semi-circle of white light. Curve B, then, represents the relative energy of filtered light at each wavelength to produce this scotopic match.[6]

Although the two curves in Fig. 6 are not precisely parallel, they serve to demonstrate that the electrical response and the scotopic luminosity curve have roughly the same peak of sensitivity in the neighborhood of 500 mμ. While the number of experimental points is small, there is some indication that Curve A for the electrical response indicates a relatively greater sensitivity to blue light than does Curve B, the scotopic luminosity curve. This is in spite of the fact that the intensities used in the electrical experiments are well above the photochromatic interval and might be expected to result in a photopic form of sensitivity curve. Such a curve would, of course, show a relatively low sensitivity to the blue and high sensitivity to the red end of the spectrum. All of our subjects have reported, indeed, that the red flashes appear very bright, even when they arouse but a small electrical response. The blue flashes, on the other hand, need not appear so bright in order to arouse large electrical responses.

[6] A more precise determination of this function, using the same procedure with a number of other subjects, is described in a forthcoming article (6). This new determination is in substantial agreement with that reported here.

Responses of the light-adapted eye to colored lights.—It was our original intention to obtain comparable data on the electrical responses of the light-adapted eye to lights of various wavelengths. In this we have been only partially successful, since it is unfortunately true that even our very intense stimuli of the longer wavelengths fail to arouse measurable responses while the eye is being flooded with light. Responses at very low levels of light-adaptation are similar to those of the dark-adapted eye, except that the response magnitudes are reduced in the same manner as shown for white light in Figs. 1 and 2. We have also failed to obtain responses for foveal stimulation alone. With the high intensities of light required to arouse a response under these conditions, there is almost certainly some scatter over the peripheral retina.

Control experiments.—A number of control experiments were performed to guard against certain obvious experimental artifacts. Deliberate winking on the part of the subject produced characteristic deflections which were easily distinguishable from true responses on the photographic records. Deliberate eye movements and changes in accommodation distance were also recorded, and were found again to have an entirely different pattern of response. The possible effect of a galvanic skin response under the neutral electrode was ruled out by the fact that no change in the response resulted when a mouth electrode was substituted.

Stimulation of the other (left) eye produced no response unless the intensity of stimulation was extremely high. Barely noticeable responses under these conditions presumably represent reflex or conditioned winks, eye movements, or pupillary responses, which are of negligible effect when superimposed upon the large retinal responses obtained when such flashes stimulate the right eye. The possibility of a shutter artifact was also ruled out by the fact that a very dim flash, or none at all, produced no deflection whatever.

In one experiment the pupillary and accommodative mechanisms were paralysed by the application of atropine. No effect was noted on the principal wave of the response. We may, therefore, conclude that the early portion of each record is not significantly influenced by the reflex activities of the intra-ocular muscles. The chief effect of atropine was to abolish the '*c*-wave' of the response. This wave normally has the form of a slow rise above the baseline, beginning as the '*b*-wave' declines and reaching a maximum about 0.5 to 1.0 sec. after the flash (see Fig. 4). Our tentative conclusion is, therefore, that this wave results from the pupillary or accommodative mechanisms.

Perhaps the most convincing evidence that the responses are truly retinal is in the fact that they resemble so closely those of the excised eyes of animals, and even those obtained by applying electrodes directly to the retina without the complicating effects of other tissues (see Kohlrausch, 4, pp. 1416–1421).

Conclusions

The data reported above are consistent with the generally accepted fact that the principal wave of the retinal response reflects primarily the scotopic system of the eye. The similarity of curves A and B in Fig. 6, obtained as they were by two very different techniques and at very different levels of stimulation, is evidence that the wavelength sensitivity of the electrical response is scotopic in character. The fact that the re-

sponse is nearly obliterated by a moderate level of light adaptation is consistent with the supposition that the scotopic mechanism has yielded to the photopic at this level. To the limited extent that responses to colored lights can be obtained in the light-adapted eye, no evidence of a Purkinje shift toward greater relative sensitivity to the longer wavelengths has been found. It seems clear that the method used in these experiments is capable of revealing the responses of the scotopic mechanism of the retina even at high intensities of stimulation. The results so obtained are not complicated by the intrusion of the photopic system or by the activity of higher visual centers.

Our records have never revealed as large an initial (presumably photopic) component as those of Adrian. Our principal *b*-wave has always been larger than this initial diphasic effect. A partial explanation of the discrepancy may lie in the fact that the resistance-capacity-coupled amplifier used by Adrian would serve to favor rapid components over slower ones, while our direct-coupled system responds equally well to all of them. It is probable, also, that individual differences among experimental subjects may account for some of the discrepancy, although later experiments have failed to reveal any very large photopic components in the responses of five other subjects (6).

Summary

1. Oscillographic records of electrical responses from the normal human retina have been obtained by a technique which yields a faithful reproduction of the magnitude and wave-form of each response.

2. When the eye is adapted to successively higher levels of field brightness both the magnitude and latency of the responses to a test flash are markedly reduced. When the level of adaptation has been raised to 1.6 millilamberts (ca. 1.5 footlamberts), the electrical responses are found to have disappeared almost completely.

3. The response to red light includes an initial diphasic wave. This is never present in responses to blue light, but may appear in reduced amount in responses to lights of intermediate wavelengths.

4. The specific sensitivity of the eye to colors (luminosity function) as revealed by this type of recording is similar to that of the dark-adapted eye as measured by the conventional low-level matching experiments.

5. The quantitative data obtained in these experiments support the conclusion that the principal component of the human retinal action potential is initiated by the scotopic system of the eye. Hence, the present technique provides a means of isolating that system for study, even at high levels of stimulation.

(Manuscript received October 22, 1948)

References

1. Adrian, E. D. The electrical response of the human eye. *J. Physiol.*, 1945, **104**, 84–104.
2. Adrian, E. D. Rod and cone components in the electric response of the eye. *J. Physiol.*, 1946, **105**, 24–37.
3. Johnson, E. P. The electrical response of the human retina during dark-adaptation. *J. exp. Psychol.*, 1949, **39**. (In press)
4. Kohlrausch, A. Elektrische Erscheinungen am Auge. In: Bethe, A., *et al.* (Ed.) *Handbuch der normalen und pathologischen Physiologie*, vol. 12, part 2. Berlin: Julius Springer, 1931, pp. 1393–1496.
5. Riggs, L. A. Continuous and reproducible records of the electrical activity of the human retina. *Proc. Soc. exp. Biol. Med.*, 1941, **48**, 204–207.
6. Riggs, L. A., Berry, R. N., & Wayner, M. A comparison of electrical and psychophysical determinations of the spectral sensitivity of the human eye. *J. opt. Soc. Amer.*, 1949, **39**, 427–436.

THE EFFECT OF INTER-SENSORY STIMULATION ON DARK ADAPTATION AND NIGHT VISION

BY A. CHAPANIS, R. O. ROUSE, AND STANLEY SCHACHTER

Departments of Psychology, The Johns Hopkins University, Williams College, and The University of Michigan [1]

Introduction

Research on the facilitation and inhibition of sensation in one modality due to stimulation of other sense organs has had a long but controversial history. Of interest in this connection are a number of recent reports from Russian laboratories (6, 7, 8, 19) which claim to show enormous effects of intersensory stimulation on dark adaptation and night vision. Thus, Kekcheyev (7) reports, "Several months ago we experimented in expediting adaptation by means of light muscular exercise. . . . Experiments made on ten subjects with the help of the adaptometer revealed that it was possible in this way to reduce the period of adaptation from 25–45 minutes to 5–6 minutes." In another article (6), the same author claims that "In some instances the period of adaptation dropped from 45 minutes to 8 minutes" with combinations of gustatory stimulants. Kekcheyev then goes on to say in the same article that auditory, olfactory, gustatory, labyrinthine, thermal, pain, tactile, proprioceptive, and interoceptive stimuli produced changes in the sensitivity of the dark-adapted eye. This is summarized by his statement, "It can be seen, therefore, that excitation of *any*[2] receptor produces changes in the sensitivity of the dark-adapted eye" (6).

Other researches show that the 'sensitizing' stimulus need not be sensed in order to change the visual threshold. Thus, it appears that ultrasonic vibrations (9), ultra-violet light (5), radio waves (21), and the geographical latitude of the laboratory where tests are made (10), affect rod sensitivity. In view of these findings, it is perhaps not surprising to find that it is also possible to demonstrate sensory conditioning with the dark-adapted eye (8). Because such sensory conditioning had presumably occurred during an individual's life-span, "the sensitivity of sight and hearing was modified not [only] [3] by the influence of physical excitors themselves (light, sound, cold, mechanical action, etc.), but only[4] by the *idea*[2] of the corresponding sensations. The changes in one direction or the other were well marked and lasted for about half an hour." Similar effects have also been reported by Lasareff and Dobrovolskaia (11) who observed that imaginary singing altered the sensitivity of the dark-adapted eye in the same way that real singing did. In this study, however, the changes lasted only a few minutes.

These reports are so insistent and the claims so extravagant that it is somewhat surprising to find virtually no experimental work to test them. In part, this situation may arise from the fact that the Russian reports are

[1] This research was carried out at the Aero Medical Laboratory, Air Materiel Command, Wright Field, Dayton, Ohio, during the fall of 1943 and spring of 1944.

[2] Authors' emphasis

[3] Authors' insert

[4] The translation appears to be faulty here. The sense of this statement is more accurately conveyed by the word 'even' instead of 'only.'

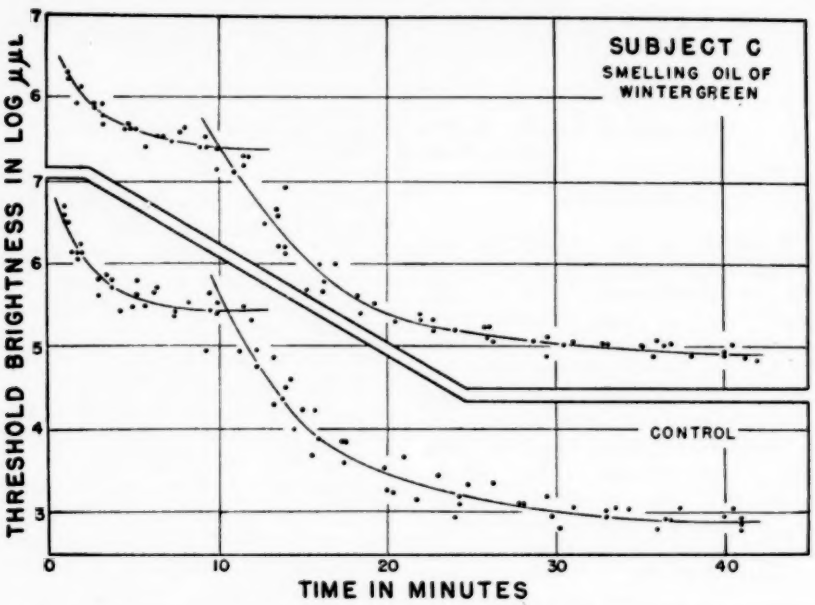

FIG. 1. The upper section contains the composite data of five series of dark-adaptation measurements for Subject C while he was smelling oil of wintergreen. The lower section contains data for a similar number of measurements obtained under control conditions.

extremely skimpy in presenting data or details of apparatus and experimentation.[5]

Studies from non-Russian laboratories contribute little material relevant to this problem. Serrat and Karwosky (17) report that auditory stimulation produced no effect on the general or specific color threshold. It is not clear from their paper whether the subjects were completely dark-adapted during these trials, although it appears likely that they were not. Thorne (20) reports a few incidental measurements on the effects of auditory stimulation on absolute visual sensitivity. His findings appear to show that the effects of auditory stimulation may be either facilitatory or inhibitory. On the basis of his own introspections and a few obtained from the subjects, he concludes that "when the simultaneous stimulus is attributively strong enough to become the figure in the 'perceptual figure-ground' relationship, it raises liminal sensitivity or exerts inhibitory effect; when it continuously occupies the ground, it facilitates with resultant lowering of threshold." Di Giorgio (2) reports that differential brightness sensitivity is increased following excitation of the vestibular mechanism. The illumination level is not specified but may be presumed to be above the cone threshold. The effect is transitory, since the differ-

[5] Attempts were made by the senior author to obtain this information in 1943 through military and diplomatic channels. These attempts met with no success. Somewhat more recently (December, 1947) the senior author attempted to communicate directly with Dr. Kekcheyev via the mails. The letter was returned unopened from Moscow.

ential threshold returns to normal within a few minutes.

The most direct experimental evidence on the validity of the Russian researches comes from experiments by Rose and Schmidt (16). They studied the effects of several factors—among others, light muscular exercise and stimulation of taste with saccharine—on dark adaptation. Neither kind of stimulus affected dark adaptation. The findings by Rose and Schmidt are at variance with results obtained by Matthews and Luczak (14) who found that light muscular exercise favored dark adaptation. The magnitude of the effect depends somewhat on which of their data are taken as the control data. In any case, the maximum possible effect amounted to 0.65 log units at 11 min. after the onset of dark adaptation and decreased thereafter. Matthews and Luczak conclude that the effects "are not nearly as great as those described by him" [Kekcheyev].[3]

In this article three experiments are reported. The first was made to determine whether smell (oil of wintergreen), sound (1000-cycle tone), or internal stimuli (light muscular exercise) have any effect on dark adaptation. The second experiment was performed to determine whether tactile or auditory stimulation has any effect on contrast sensitivity. The third tested the effect of similar kinds of stimuli on form discrimination at very low brightness levels. These experiments were done in 1943 and early 1944, but the results were so

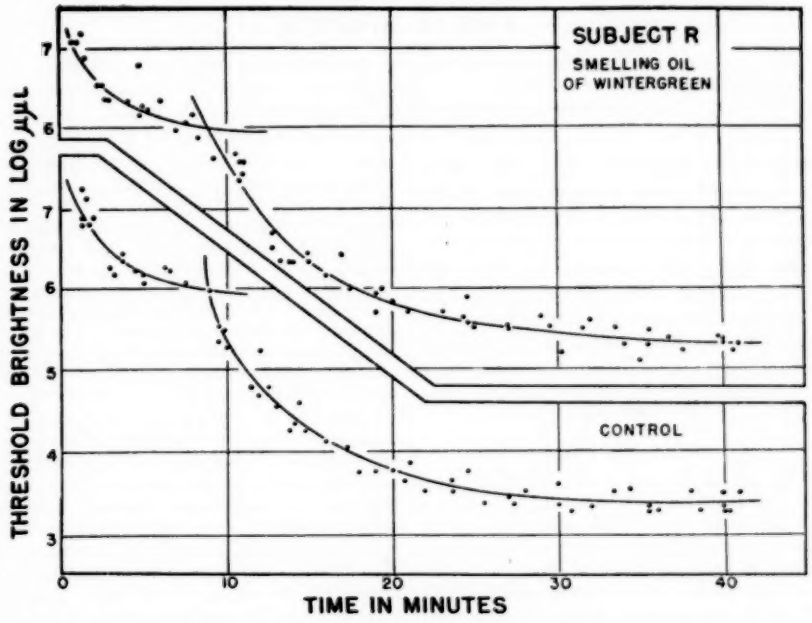

FIG. 2. The upper section contains the composite data of five series of dark-adaptation measurements for Subject R while he was smelling oil of wintergreen. The lower section contains data for a similar number of measurements obtained under control conditions.

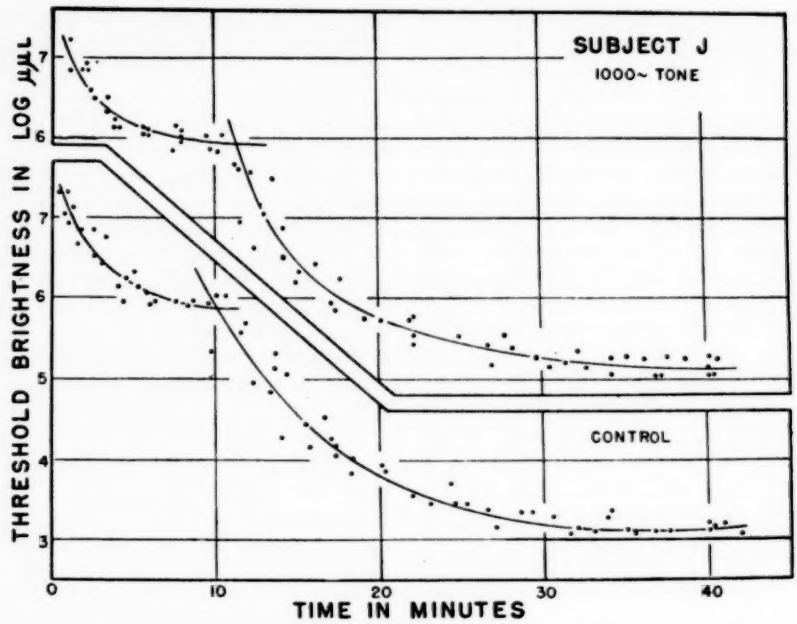

FIG. 3. The upper section contains the composite data of five series of dark-adaptation measurements for Subject J while he was listening to a 1000-cycle tone. The lower section contains data for a similar number of measurements obtained under control conditions.

uniformly negative that the data have never been summarized or reported. In view of the paucity of experimental literature on these problems and the continuing flow of articles from Russian sources, the negative findings obtained in these experiments may be of some general interest.

EXPERIMENT I

Apparatus and Procedures

The apparatus used in this experiment was the large Hecht-Shlaer adaptometer (4). A pre-adaptation period of three min. preceded all measurements. During the pre-adaptation period, *S*s fixated a white field about 35° in size with a brightness of about 1800 millilamberts. Threshold measurements were made with a 3° violet test light situated 7° nasally from a fixation light in the visual field. The stimulus light was interrupted for one-half sec. every sec. by a motor-driven, segmented disc in the light path of the adaptometer. Measurements were made about every two or three min. during the entire course of dark adaptation. *S*s were allowed to determine their own thresholds by adjusting the optical wedge until they could 'just see the flickering light.' The knob which controls the wedge from the *S*'s side of the apparatus is conveniently located near the eyepiece. This technique differs from the usual procedures for determining thresholds with this instrument. It was used to avoid any possibility that the data might be biased by *E*. During the initial stages of dark adaptation, the threshold changes so rapidly that it is impossible to use any complete psychophysical procedure because of the time required. In usual practice, therefore, *E* takes a rapid series of determinations with an ascending and descending series of brightnesses and then estimates the threshold on the basis of the pattern of *S*'s responses. In this procedure there is at least the possibility that *E* might unconsciously bias the data in one direction or the other. Since the *S*s had no knowledge of their results, and only minimum knowledge of the

purpose of the experiment, it seemed probable that they would be much less likely to introduce any bias into the data with the modified procedure used in this study. The use of the segmented disc also had the advantage that determinations of the threshold were more rapid than when ordinary methods were used. Thus transitory changes would be more likely to be recorded.

Six *Ss* were used in this experiment. Two were subjected to an experimental regime with the smell stimulus, two with the sound stimulus, and two with the muscular exercise. Each *S* yielded a total of 11 dark-adaptation curves on as many different days. The first day's dark-adaptation measurements were used to familiarize the *S* with the apparatus and experimental procedures. The results obtained on this day were not used in the subsequent tabulation of data. The experimental and control tests were given in a balanced order to each *S* on the 10 succeeding days. One *S* in each pair was assigned the order: T C C T C T T C C T; and the other: C T T C T C C T T C, where the T refers to the test or experimental day and the C refers to the control day. In this, and the two experiments described below, *Ss* were always tested individually.

The 'sensitizing' stimuli used in this experiment cannot, unfortunately, be described quantitatively. In the olfactory experiments, the *Ss* inhaled the odors emanating from a small bottle containing oil of wintergreen. This odor is used for testing leaks in oxygen masks and is rather strong and pungent. It is not offensive, however, and cannot be described as overpowering. The auditory stimulus, a 1000-cycle tone of about 50 db. intensity, was produced by a Hewlett-Packard oscillator. For the muscular exercise the *Ss* were required to ride a bicycle ergometer. The exercise can best be described as being very light. None of the *Ss* reported fatigue after performing this exercise.

In the experimental runs the *Ss* were exposed to the pre-adapting light, as is customary. Immediately following the preadaptation period, *Ss* were exposed to the 'sensitizing' stimulus, which was continuously applied throughout the

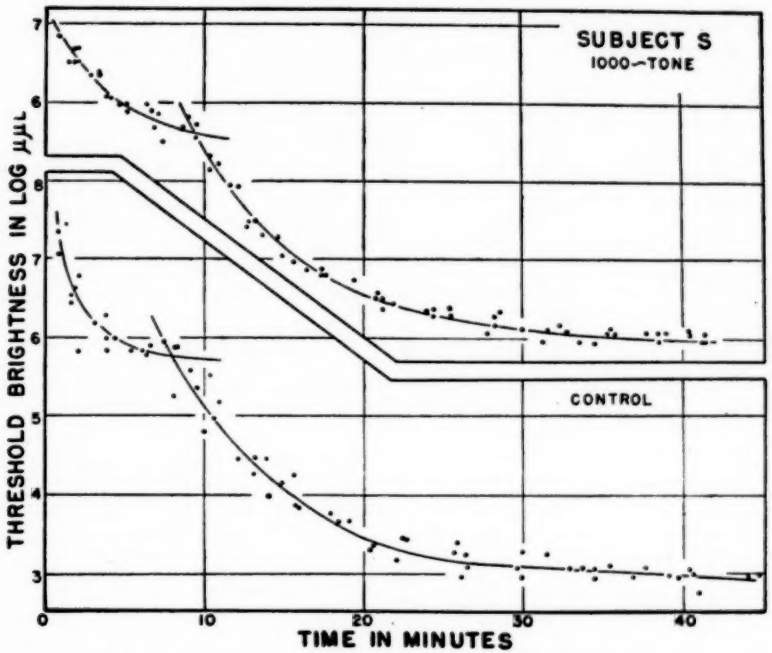

FIG. 4. The upper section contains the composite data of five series of dark-adaptation measurements for Subject S while he was listening to a 1000-cycle tone. The lower section contains data for a similar number of measurements obtained under control conditions.

entire course of dark adaptation. In the trials involving the light muscular exercise, *Ss* mounted the bicycle ergometer, rode it for about three min., dismounted, and made a threshold measurement. Following this measurement they again mounted the bicycle and continued exercising for about three min. until the next measurement was taken. Because of the intermittent nature of this stimulus situation, the data obtained in the exercise trials are considerably fewer than in the control runs, and the threshold readings tend to occur at fixed time intervals

Results

The results of this experiment are shown in Figs. 1 through 6. There are a total of five dark-adaptation curves contributing to each set of control data and five dark-adaptation curves contributing to the data for each experimental condition. There is no evidence that any of the stimuli used in this experiment either facilitated or inhibited dark adaptation. The thresholds vary somewhat from day to day, as is to be expected, and differences between the experimental and control runs are well within this normal range of variation.

Experiment II

The second experiment was designed to determine whether tactile and auditory stimuli affect low contrast sensitivity.

Apparatus and Procedures

The apparatus used in this experiment was the Luckiesh-Moss Low-Contrast Test-Chart (13), illuminated to a brightness of 0.014 apparent foot candles. *Ss* read the chart from a

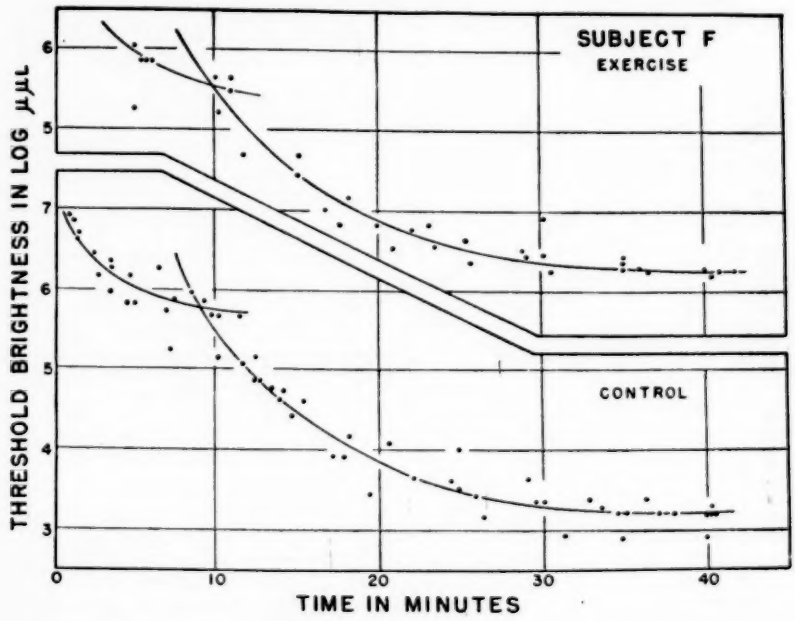

Fig. 5. The upper section contains the composite data of five series of dark-adaptation measurements for Subject F. Threshold determinations were interspersed with three-min. periods of light muscular exercise. The lower section contains data for a similar number of measurements obtained under control conditions.

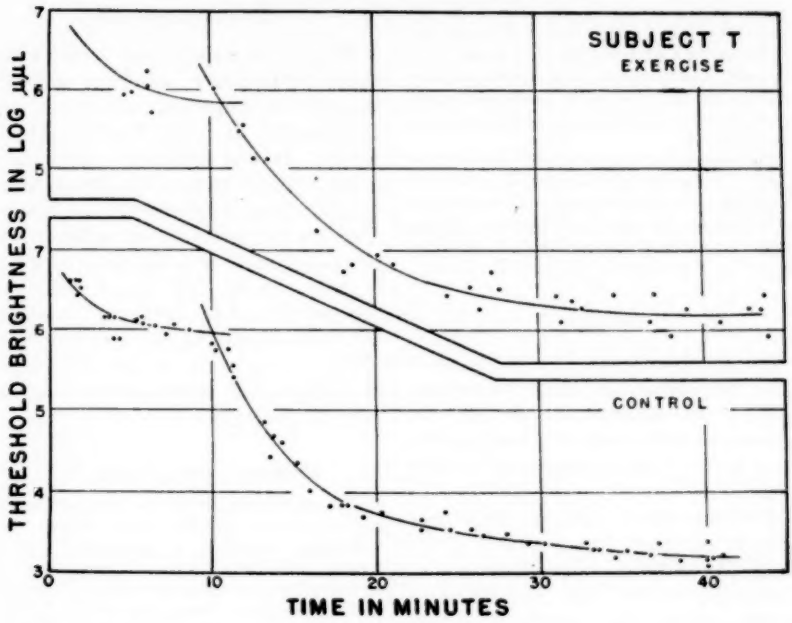

FIG. 6. The upper section contains the composite data of five series of dark-adaptation measurements for Subject T. Threshold determinations were interspersed with three-min. periods of light muscular exercise. The lower section contains data for a similar number of measurements obtained under control conditions.

distance of 10 feet after a dark-adaptation period of 15 min. Coincident with their reading of the chart, *Ss* were stimulated with either of two intensities of pressure or two intensities of sound. The heavy pressure was produced by a

TABLE I

EXPERIMENTAL DESIGN USED IN THE STUDY OF THE EFFECT OF INTERSENSORY STIMULATION ON CONTRAST SENSITIVITY

The entries are the days on which the experimental conditions were presented.

Subjects	Experimental Conditions				
	Loud Sound	Weak Sound	Heavy Pressure	Light Pressure	Control
A	2	3	4	5	1
B	4	1	5	2	3
C	1	5	2	3	4
D	5	4	3	1	2
E	3	2	1	4	5

TABLE II

NUMBER OF CORRECT RESPONSES MADE BY FIVE SUBJECTS IN READING THE LUCKIESH-MOSS LOW CONTRAST TEST CHART UNDER VARIOUS EXPERIMENTAL CONDITIONS

Subjects	Experimental Conditions					Means
	Loud Sound	Weak Sound	Heavy Pressure	Light Pressure	Control	
A	21	22	20	22	22	21.4
B	22	16	23	19	23	20.6
C	14	14	23	24	20	19.0
D	29	24	24	24	28	25.8
E	16	15	14	15	13	14.6
Means	20.4	18.2	20.8	20.8	21.2	
	Experimental Days					
	1	2	3	4	5	
Means	18.0	21.2	21.8	20.2	20.2	

500-gm. weight applied to the back of the *S*'s hand. The light pressure consisted of a 40-gm. weight similarly placed. The loud sound had a frequency of 3800 cycles per sec. and an intensity of about 70 db. The weak sound had a frequency of 2800 cycles per sec. and an intensity of about 30 db.

The experiment was designed as a Latin square. Table I exhibits the experimental routine and is self-explanatory. *S*s were tested on five consecutive days, and the order of presentation of the various experimental conditions was varied systematically from day to day.

Results

The results obtained in this experiment are shown in Table II, which reports the number of correct responses made by each *S* in reading the chart under each experimental condition.

An analysis of variance was performed on the data presented in Table II, with the results shown in Table III. From this analysis it is evident that the only significant source of variation was that between *S*s. There were no consistent variations from day to day, and the variance which can be attributed to the different experimental conditions is *less* than the error variance. There can be little doubt, therefore, that the introduction of these stimuli into the *S*s' environment had no effect on their contrast thresholds.

TABLE III

ANALYSIS OF VARIANCE OF THE DATA IN TABLE II

Source of Variance	Sums of Squares	Degrees of Freedom	Estimated Population Variance
Between Subjects	328.64	4	82.16*
Between Days	41.84	4	10.46
Between Experimental Conditions	28.64	4	7.16
Residual	91.92	12	7.66

* Significant at the one percent level

EXPERIMENT III

The third experiment was undertaken to measure the effect of intersensory stimulation on form discrimination at very low brightness levels. It was originally planned as a large Latin square involving seven *S*s and seven experimental conditions. Due to the pressure of other work, the experiment was terminated when it was about half completed. Although complete data are available for only three *S*s, it seems worthwhile to report them here since they contribute additional evidence on this general problem.

Apparatus and Procedures

The apparatus used was the Johnson Foundation Luminous Plaque Adaptometer. This instrument contained a luminous T which could be oriented in any one of four positions by *E*. At a distance of six feet—the test distance—the T subtended a visual angle of 3°. Crossed polaroid discs permitted adjustment of the brightness of the test object from 4.2 to 2.8 log $\mu\mu$L in steps of 0.2 log $\mu\mu$L. A shutter obscured the test object between trials.

*S*s were dark-adapted for 30 min. by wearing red goggles in ordinary room illumination and for an additional 10 min. by sitting in total darkness. Ten presentations of the test object were made at each brightness level, starting with the highest. Trials were continued at successively lower brightness levels until the *S* missed six or more presentations at any level. The score is recorded simply as the number of trials in which the position of the T was correctly reported. No fixation point was provided for the *S*s; no time limit was placed on individual trials; and *S*s were forced to guess whenever they were uncertain.

The design of that portion of the experiment which was completed is shown in Table IV. Tests were made on eight separate days for each *S*. The first day was devoted to a lecture on night vision, demonstrations, and a practice test. Results obtained on this day were not recorded. Three of the stimuli used (oil of wintergreen, 1000-cycle tone, and exercise), were identical to those used in the first experiment. The new smell stimulus, oil of eugenol, was introduced by having the subjects inhale the odor from a small vial containing about three cc. of the liquid. Eugenol is derived from cloves

TABLE IV

EXPERIMENTAL DESIGN USED IN THE STUDY OF THE EFFECTS OF INTER-SENSORY STIMULATION ON FORM DISCRIMINATION AT LOW BRIGHTNESSES

Entries are the days on which each condition was tested.

Subjects	Experimental Conditions						
	Control	Smell Wintergreen	Smell Eugenol	Taste Lime	1000~ Tone	Noise Buzzer	Exercise
A	3	2	5	4	1	7	6
B	2	5	7	1	4	6	3
C	7	6	1	2	5	3	4

TABLE V

NUMBER OF TEST OBJECTS CORRECTLY SEEN ON THE JOHNSON FOUNDATION ADAPTOMETER UNDER VARIOUS EXPERIMENTAL CONDITIONS

Subjects	Experimental Conditions							
	Control	Smell Wintergreen	Smell Eugenol	Taste Lime	1000~ Tone	Noise Buzzer	Exercise	Means
A	61	60	58	63	60	64	57	60.4
B	45	55	53	48	46	37	52	48.0
C	47	53	49	51	46	53	51	50.0
Means	51.0	56.0	53.3	54.0	50.7	51.3	53.3	

or cinnamon and has a fairly strong but agreeable spicy aroma. The taste stimulus was administered by having *S*s suck half a thin slice of fresh lime. The buzzer had an intensity of about 50 db. Stimuli were initiated five min. before the actual tests began. On the exercise trials, *S*s were also instructed to squeeze a hand dynamometer gently between trials.

Results

The results of this experiment are shown in Table V and the analysis of variance of the data is shown in Table VI. Because of the incomplete nature of these data, the analysis may perhaps be suspect. Two alternative analyses were also computed: (1) by extracting 'between subjects' and 'between days' variance, and (2) by extracting 'between subjects,' 'between days,' and 'between experimental conditions' variance. No method of analysis yielded a 'between days' or 'between conditions' variance which was even as large as the residual variance. There is not the slightest evidence, therefore, that the experimental conditions had any effect on this type of visual performance.

TABLE VI

ANALYSIS OF VARIANCE OF THE DATA OF TABLE V

Source of Variance	Sums of Squares	Degrees of Freedom	Estimated Population Variance
Between Subjects	623.5	2	311.8*
Between Treatments	66.6	6	11.1
Residual	241.1	12	20.1

* Significant at the one percent level

DISCUSSION

Before performing the experiments reported here, we obtained about 20 dark-adaptation curves under various

experimental conditions. These were not collected systematically and were in the nature of exploratory experiments. A variety of stimuli—smells, tastes, noises, pure tones, and exercise—was used in these preliminary trials. It is important to note here only that there was not the slightest evidence in any of these trials that intersensory stimulation affected dark adaptation.

One difficulty with experiments of this sort is that the human body is always being stimulated, so that it is difficult to define precisely what constitutes a 'control.' Although attempts were made to keep the *S*s as quiet and relaxed as possible during the control tests, the *E*s were not always successful. For example, all of these experiments were carried out in a dark room which had fairly good acoustical insulation. On two occasions, however, loud noises due to an air hammer being used on nearby construction occurred during control runs. As another example, *S*s sat in a swivel chair during the dark-adaptation experiments. On one occasion, *E* discovered that the *S* had been vigorously swinging back and forth between threshold determinations. On still another occasion, one *S* reported extreme internal discomfort due to colonic distention. Fits of coughing and sneezing were occasionally noted and still other external and internal stimuli were undoubtedly impinging almost constantly on the *S*s. The control sessions were designed to involve a minimum of extraneous stimuli, but these examples indicate that it is impossible even to duplicate the stimulus situation from trial to trial. The data from these disturbed control sessions were included as control data, since they do not differ from data in more placid periods.

Some of the discrepancy between the findings of this study and those of Kekcheyev *et al.* can perhaps be attributed to the way in which the two groups of data are reported. As has already been indicated, the data in the Russian studies are extremely skimpy, but those data which are given are reported in terms of retinal 'sensitivity,' i.e., *E* (Empfindlichkeit). *E* was used by Aubert, Piper, Nagel, and other early German investigators in this field. It is defined as the reciprocal of the threshold brightness of the light source. In his early work on dark adaptation, Hecht (3) expressed dissatisfaction with this measure because (a) it did not lend itself to easy interpretation in theoretical analyses of the visual process, and (b) it greatly distorted the real nature of the dark adaptation process. The nature of this distortion is illustrated in Fig. 7. The upper chart shows a dark adaptation curve (solid line) in which threshold sensitivity is plotted along the ordinate in terms of the logarithm of the brightness of the just perceptible light. One advantage of a logarithmic plot is, of course, that equal percentage changes are represented by equal linear distances on the chart.

It is generally agreed by workers in this field that dark adaptation measurements made during any one experimental session may easily vary in a random manner as much as 0.1 log units from the ideal curve. As a matter of fact, when continuous recording methods have been used for measuring dark adaptation, variations much greater than this have been observed (1). Deviations of ±0.1 log units are shown by the interrupted lines in the upper chart. Compared with the total range of dark adaptation, this amount of variation appears quite small in the logarithmic plot.

The lower chart in Fig. 7 shows the

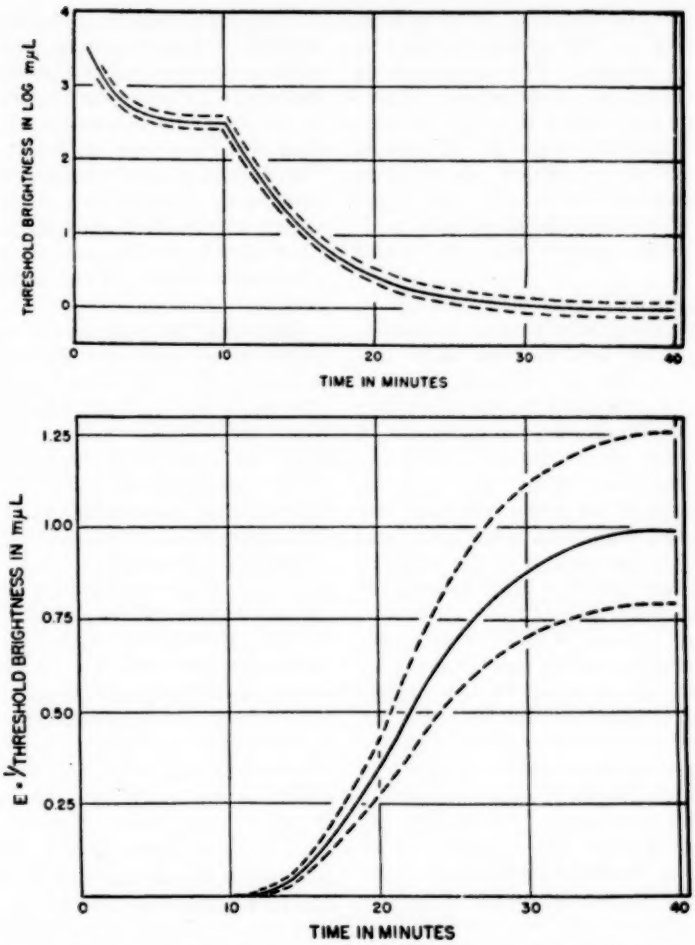

FIG. 7. The upper section shows an average dark-adaptation curve (solid line) with the normal limits of random fluctuations which can be expected during any one session (interrupted line). The ordinate scale is threshold brightness in log milli-microlamberts. The lower section shows the same data plotted in terms of sensitivity, i.e., the reciprocal of threshold brightness. Note the distortion introduced by the latter method of plotting these data.

same data plotted in terms of *E*, defined as the reciprocal of the threshold brightness. The distortion introduced by this method of graphing the data is immediately apparent. Changes in the sensitivity of the eye during the first 10 min. of dark adaptation—when expressed in *E* units—are so small that they cannot be graphed. Yet at 10 min. the eye can see a light which is one-tenth as bright as a light which can just be seen at one min. It is much more important for our purposes, however,

to notice the distortion introduced in the data for the terminal stages of dark adaptation. The normal 0.1 log unit variations in threshold (shown by the interrupted lines) are greatly magnified in the plot of *E*. Expressed in terms of percentages—as many of the Russian results are—the upper limit of this variability at 40 min. is 60 percent greater than the lower limit.

Daily variations and differences between individuals are, of course, considerably greater than those discussed above. In Piper's classical data (15), for example, the difference between the lowest and highest *E* values for his 17 *S*s is enormous—amounting to about 592 percent. Expressed in log units, however, this is a range of about 0.79 log units, and is roughly comparable to the range of thresholds found in large samples of normal individuals (18).

Many of the Russian results are expressed as percentage changes in *E* from the control to the experimental conditions. In the study by Lasareff and Dobrovolskaia (11), for example, singing, real or imaginary, produced changes in *E* which amount to about 40 percent. In one of Kekcheyev's studies (6), smelling oil of bergamot increased sensitivity by 15 percent. In another article (5) which summarizes a number of researches, the *greatest* effect reported for any *S*, with *any* kind of simulus, was an increase in *E* from 1.69 to 7.70. This is an increase of 0.65 log units. It is impossible to determine from any of the Russian studies precisely how much variation in *E* was encountered in any control run. From the analysis presented here, however, it is at least plausible that most of the changes they report fall within the range of intra-experimental variation. It is barely possible that random fluctuations in threshold,[8] greatly magnified by their expression in terms of *E*, might have been misinterpreted as real effects. Certainly, no significant or consistent improvement in dark adaptation could be demonstrated if our data were reported in terms of *E*. The fact that *any* stimulus—real or imaginary, adequate or non-adequate—seems to produce changes in visual sensitivity in the Russian laboratories also makes their results highly suspect, in our opinion.

Although this criticism and analysis applies to many of the Russian findings, it cannot explain the decrease in *time* required for dark adaptation—from 30 min. to 5–8 min. (6, 7). Since the data supporting this claim have apparently never been published, it is impossible to criticize this aspect of their findings. The experiments reported in this paper, however, lend no support to this claim. It may well be that variations in the intensity of stimuli produce facilitatory or inhibitory effects, as has been claimed by Thorne. If this is the case, the results are probably so small that extremely precise techniques must be used to demonstrate them. From a practical standpoint, therefore, there does not appear to be any basis for supposing that the dark adaptation of military men in combat can be significantly altered by intersensory stimulation or exercise.

Summary and Conclusions

Three experiments were performed to test recent Russian claims that intersensory stimulation and light

[8] Lee *et al.* (12), for example, report wave-like fluctuations in dark adaptation records with a three to seven minute period and with amplitudes of 0.05 to 0.25 log units. The fact that these periodic fluctuations in one eye are not necessarily correlated with the fluctuations in the other eye makes it highly improbable that they are produced by the stimulation of other sense organs.

muscular exercise significantly improve dark adaptation. In the first experiment, six *Ss* were tested with a Hecht-Shlaer adaptometer on 11 consecutive days. Experimental treatments consisted of smelling oil of wintergreen, listening to a 1000-cycle tone, and performing light muscular exercise on a bicycle ergometer. Control and experimental dark-adaptation sessions were varied systematically from day to day. In the second experiment, five *Ss* were tested on five consecutive days with the Luckiesh-Moss Low-Contrast Test-Chart. The stimuli used here were a loud and soft tone and a heavy and light pressure applied to the back of the hand. The third experiment measured the effects of similar stimuli on form discrimination at very low illumination levels. The results of all experiments are completely negative. None of the stimuli used in this experiment either facilitated or inhibited dark adaptation, contrast sensitivity, or form discrimination at low illuminations.

(Manuscript received August 9, 1948)

References

1. Craik, K. J. W., & Vernon, M. D. Perception during dark adaptation. *Brit. J. Psychol.*, 1942, **32**, 206–230.
2. Di Giorgio, A. M. Modificazioni del punto prossimo, dell'acuità visiva e del potere di discriminazione dell'intensità luminosa, in seguito ad eccitamento del labirinto. *Boll. Soc. ital. Biol. sper.*, 1940, **15**, 391–394.
3. Hecht, S. The dark adaptation of the human eye. *J. gen. Physiol.*, 1920, **2**, 499–517.
4. Hecht, S., & Shlaer, S. An adaptometer for measuring human dark adaptation. *J. opt. Soc. Amer.*, 1938, **28**, 269–275.
5. Kekčeev,[7] K. Ch. On the action of non-adequate stimuli on receptors. *C. R. Acad. Sci. U. R. S. S.*, 1937, **14**, 495–497.
6. Kekčheev,[7] K. K. Methods of accelerating dark adaptation and improving night vision. *War Med.*, 1945, **8**, 209–220.
7. Kekcheyev,[7] K. Expediting visual adaptation to darkness. *Nature*, 1943, **151**, 617–618.
8. Kekcheyev,[7] K. Conditioned excitors and human sense organs. *Nature*, 1945, **156**, 573–574.
9. Kektchéeff,[7] K. C., & Ostrovsky, E. P. Sur la définition des vibrations aériennes de fréquence ultra-sonore par la mesure des seuils optiques. *C. R. Acad. Sci. U. R. S. S.*, 1941, **31**, 370–372.
10. Lasareff, P. P. Théorie thermique des changements de la sensibilité visuelle périphérique produits par les causes géophysiques. *C. R. Acad. Sci. U. R. S. S.*, 1937, **14**, 275–278.
11. Lasareff, P. P., & Dobrovolskaia, E. V. De l'influence du chant sur l'adaptation au cours de la vision périphérique. *C. R. Acad. Sci. U. R. S. S.*, 1937, **16**, 271–272.
12. Lee, R. H., Finch, E. M., & Pounds, G. A. Periodic fluctuations in the dark adapted threshold. *Amer. J. Physiol.*, 1945, **143**, 6–10.
13. Luckiesh, M., & Moss, F. K. The ability to see low contrasts at night. *J. aero. Sci.*, 1942, **9**, 261–263.
14. Matthews, B. H. C., & Luczak, A. K. Some factors influencing dark adaptation. Flying Personnel Research Committee Report No. 577, RAF Physiological Laboratory, Farnborough, June, 1944. 33 pp.
15. Piper, H. Über Dunkeladaptation. *Z. Psychol. Physiol. Sinnesorg.*, 1903, **31**, 161–214.
16. Rose, H. W., & Schmidt, I. Factors affecting dark adaptation. *J. Aviat. Med.*, 1947, **18**, 218–230.
17. Serrat, W. D., & Karwoski, T. An investigation of the effect of auditory stimulation on visual sensitivity. *J. exp. Psychol.*, 1936, **19**, 604–611.
18. Sheard, C. Dark adaptation: some physical, physiological, clinical, and aeromedical considerations. *J. opt. Soc. Amer.*, 1944, **34**, 464–508.
19. Streltsov, V. The function of the eye in aviation. *Amer. Rev. Sov. Med.*, 1944, **2**, 127–133.
20. Thorne, F. C. The psychophysical measurement of the temporal course of visual sensitivity. *Arch. Psychol., N. Y.*, 1934, No. 170, 66 pp.
21. Turlygin, S. J. Effect of electromagnetic centimetre waves on the central nervous system. *C. R. Acad. Sci. U. R. S. S.*, 1937, **17**, 19–21.

[7] Kekcheyev's name has been translated in many ways. The spellings given here are as they appear in these various articles; they all refer to the same author.

A SCALE OF SUBJECTIVE BRIGHTNESS *

BY R. M. HANES **

The Johns Hopkins University

The problem of measurement in psychology is an old and involved one, and one of which psychologists are well aware. Physicists and some psychologists have either questioned or even emphatically denied that the term measurement can be applied to psychological scaling. Although no attempt will be made in the present experimental report to discuss or review the theoretical aspects of scaling, it should be noted that the matter is by no means settled. While much time and effort have been devoted to the problem, there is still very fundamental disagreement.

Even if the scaling of psychological magnitudes should eventually prove to be an unsatisfactory example of measurement from the standpoint of physics, there is still, of course, both practical and theoretical value in establishing functional relationships between psychological and physical magnitudes. The success met with in constructing scales by means of the fractionation technique for attributes of auditory sensation (**4, 5**) led to the present effort to do the same thing for brightness. The sensation of brightness is affected by so many factors that there might exist considerable doubt as to the feasibility of using a method like fractionation with any promise of useful results. The aim of the present experiment was to determine how easily and how reliably individuals can make such fractional estimates of brightness in a prescribed situation.

Once it has been determined that a procedure such as this can be satisfactorily employed in constructing a brightness scale, then there is opened for investigation all of the multitude of variables which might affect the generality of such a scale and limit its usefulness. The size, spectral quality, and surround of the stimulus, the retinal location and state of adaptation, and the experience and attitude of the observer might be mentioned among these variables. Only a very few of such possible factors were studied in the present experiment. The establishment of a working scale around which the variables can be systematically investigated would logically come first

Apparatus

Since one of the goals of the experiment was the obtaining of data which could be given practical application in radar work, a stimulus situation similar to that with which a radar operator is presented was selected for study. The actual type of radar presentation commonly used in this laboratory was not chosen, because of the difficulty in obtaining physical brightness measurements on this type of display (**2, 7**). The basic difference between this latter type of presentation and the type used in this experiment lies in the manner in which the cathode-ray tube's electron beam is deflected, i.e., the scanning method. The commonly used type of radar referred to makes use of a radial scan, while the 'raster' presentation used here employs a simultaneous vertical and horizontal scan. In other words, the raster is a square pattern of light on the tube face due to the excitation of the phosphor by the simultaneous

* This research was carried out under Contract N5ori-166, Task Order I, between the Special Devices Center, Office of Naval Research, and The Johns Hopkins University. This article is Report No. 166-I-51, Project Designation No. NR-784-001, under that contract.

** The author is indebted to Dr. S. B. Williams, who supervised the writing of the manuscript, and to Dr. N. R. Bartlett, Dr. C. T. Morgan, and Dr. W. R. Garner, who read the manuscript and made helpful comments.

horizontal and vertical deflection of the tube's electron beam. (This is similar to television presentation and is commonly used in work with cathode-ray tubes.) The raster display provides a uniform light distribution not found in the radial scan and therefore permits better control of brightness, while giving essentially the same quality of light. For the instrument used in this study the horizontal rate of scan was one kilocycle and the vertical rate was 7.4 megacycles; i.e., the beam moved across the tube face one thousand times a second and at the same time oscillated up and down 7.4 million times per sec.

The brightness of this raster pattern was controlled by the amount of beam current through the instrument. Since the instrument controls read only in beam current and volts bias on the tube, a calibration in physical brightness units was necessary. This calibration was accomplished by means of a Macbeth Illuminometer. The final working calibration was the result of the readings of two *O*s. Each *O* took 10 readings at each of eight bias or beam current levels set on the instrument; the means of each of these sets of 20 readings were then calculated and plotted against bias. A smooth curve was fitted by eye to the plotted points. Since the two *O*s differed very little in their readings, their results were combined and the means taken as the best estimate of the actual physical brightness of the test field.

The square raster pattern was variable in size, and for a given beam current the brightness was, therefore, also somewhat dependent upon size. (Brightness varies inversely as size in this cathode-ray tube, since more energy strikes a unit area in a unit of time for a given rate of scan.) A one-in. square was used for this experiment. This particular size was chosen so that a relatively high level of brightness would be obtained without having the problem of the effect of size upon intensity entering to complicate the results. A much smaller (and, therefore, maximally brighter) square would have given a greater brightness range, but the results obtained from such a limited size would not be as generally applicable to a practical situation.

Because the edges of the raster pattern (especially the upper and lower ones) were noticeably brighter than the rest of the square, a black screen containing a ¾-in. hole was placed over the tube face. Thus, *O* saw nothing but a disc of light, quite uniform in brightness over its entire area. This test patch subtended an angle of about 3½ degrees at the viewing distance of 12 in.

The raster could be made to give a steady light or to alternate light and dark at various rates. The repetition rate used for most of this experiment was about 23 cycles per min., that is, the standard was on for about 1.3 sec., followed by the variable for about 1.3 sec., followed by the standard, etc. This was the maximum duration of flash which could be obtained with the apparatus. Once the first or light phase of the cycle was set at a given brightness (zero to about 100 millilamberts), the second or dark phase could be varied continuously through any brightness in the range from equality with the standard to zero brightness.

Over the face of the tube there was fastened a special hood, a funnel-shaped affair made of a stiff black felt-like material reflecting practically no light. The outer end of this hood is formed to fit around the eyes of *O* in the manner of a pair of goggles. Thus, *O*, when looking into the hood, could see nothing but the test spot on the tube; he was instructed to look directly at this spot in making his estimates.

A two-mm. artificial pupil was employed for viewing with the right eye; an opaque shield prevented light from striking the left eye. There was also an arrangement in the hood for holding Wratten neutral and color filters. This device could hold two filters at the same time. The neutral filters were used to obtain the various brightness levels for the standard; the brightness of the variable flash was altered by means of the instrument controls once the standard brightness level was set.

The cathode-ray tube used for this work was provided with a P-4 phosphor. This type of phosphor gives a white light (essentially unchanged in spectral quality regardless of intensity) with a very slightly bluish tint and has a very rapid rate of excitation and decay.[1] A short decay time was obviously essential in the type of experiment performed. Phosphors having long decay times would not be at all suitable because of the varying brightnesses due to decay.

[1] The Allen B. DuMont Laboratories manufacture the cathode-ray tube used and designate it as a 5JP4A. They describe the phosphor as having a color composition chosen in such a way that even for long observation a minimum of fatigue will be caused. A representative spectral composition curve for stationary spot excitation presented by DuMont shows peaks at about 460 mμ and 560 mμ. This distribution of energy gives a light which "appears white to the eye."

The P4 phosphor is described as having 'medium persistence.' The blue component actually decays to about one percent in five msec., while the yellow requires only one msec. to reach the same value.

Procedure

As was stated before, we were concerned in this study with determining the relation of subjective brightness or brilliance to the physical brightness of the stimulus. The fractionation technique was chosen as the method best suited both to give an answer to the practical problem involved and to allow the construction of a subjective brightness scale. *O* was required to adjust the variable stimulus, i.e., the second or dark phase of the light-dark cycle mentioned earlier, until he was satisfied that the variable was ⅓ or ¾ as bright as the standard. Since the arrangement of the apparatus did not permit *O* to manipulate the controls, it was *E* who actually made the adjustment, at the direction of *O*. This arrangement works quite satisfactorily after a little practice. *O* soon develops a specialized vocabulary for directing *E*, making use of such words as 'shade,' 'hair,' etc., to indicate the degree of brightness change that he wishes.

None of the *O*s (there was a total of seven at various times) expressed any doubts about being able to make the fractional judgment of brightness. The nature of the judgment did not seem to be at all unnatural or artificial to them. The consistency with which the judgments were made would seem to support this conclusion.

While no formal instructions were read to the *O*s, each one was acquainted with the nature of the task before the judgments were made. In addition, the main results which are to be presented were those of four individuals who had been given a fair amount of practice in the same situation previously. It was suggested to the *O*s that they use darkness as a reference or zero point for making their fractional estimates. The black disc surrounding the test spot served well for this purpose. Each one was also told that the second or darker flash was the variable, even though this fact was obvious from the way in which the second flash was varied. No time limit was placed upon any judgment; the *O*s were told to take their time and make as accurate an estimate as possible.

All observations were made in a darkened room which was almost completely light tight. In any case, a small amount of room light would not have made any difference, as shown by preliminary data. The darkened room was used principally with a view to the reduction of variability. No attempt was made, however, to maintain any given level of light- or dark-adaptation. Preliminary work had shown that there was no significant change in the judgments from the beginning to the end of an experimental period (one hour) as there should have been if the level of adaptation had any noticeable effect. Of course, the range of stimuli used was quite limited and very high brightness values might yield different data.[3] However, since the judgments in this case are of relative brightness, it seems quite probable that a very great difference in brightness levels would be needed before any marked effect could be obtained. In other words, it would seem likely that if the eye were adapted to a very high brightness, for example, there would be a nearly proportional decrease in brightness of two comparison stimuli of, say, 100 and 10 millilamberts so that the fractional judgment setting would remain the same.

The variable was set before each trial at a brightness either close to that of the standard or at a very low level, considerably below that which could possibly be called 'half.' These brightness settings of the variable were varied in a random fashion. The standard brightnesses were presented in a systematic way, either the lowest or the highest brightness being given first and successively higher or lower brightnesses being presented in order. The starting brightness level was alternated between high and low on successive trials and on successive days.

The specific procedure for each experimental period was then as follows. As soon as *O* had entered the experimental room and was seated facing the test field, the room lights were turned out. Within a minute or two the fractionation procedure was started. *E* set the standard by placing the proper neutral filter in position (no filter was used with the highest standard brightness) and *O* directed *E* to adjust the variable until he (*O*) was satisfied that the variable was ⅓ (or ¾) as bright as the standard. As soon as this judgment was made *E* set the next higher (or lower) standard brightness and the procedure was repeated. Each experimental period lasted for approximately one hour. The *O*s were given rest intervals during this hour if they so desired, but there were very few times when such intervals were requested.

Experimental Design

The principal factors investigated can be divided roughly into three parts: (a) estimates of ⅓ and ¾ brightnesses at a constant duration of flash of 1.3 sec. for both the standard and the variable; (b) the effect of flash duration on the fractionation judgments; and (c) the effect of hue on the fractionation judgments.

(a) The first part of the experiment was undertaken to answer the question: how con-

[3] The influence of surround and adaptation level, among other factors mentioned earlier in the paper, are to be investigated by the author in later experiments.

sistently and how accurately can one make fractional estimates of brightness? The ¼ estimates were included for the purpose of checking the accuracy of the half judgments and also to supply some information about the fineness of fractional estimates which could be made. Four sets of half brightness estimates were taken at each of the eight brightness levels used for each of five days for each *O*. The same number was obtained for the ¼ brightness estimates. These two types of fractional estimates were taken on alternate days. Thus, a total of 20 half judgments and 20 ¼ judgments was obtained for each individual at each brightness level.

(b) In order to be sure that time was not a possible variable influencing the brightness estimates, the second part of the study was run. Two different time patterns were used. In the first, both the standard flash and the variable flash had a duration of about 0.5 sec. Only four brightness levels were used for this part, just enough to provide data for checking the significance of any differences. For the second time pattern, the standard was on for about 30 msec. while the variable had a duration of about 970 msec. Seven standard brightnesses were employed here. For each of these time patterns at each brightness level ten estimates were made by each *O*.

(c) The last part of the experiment was designed to determine how well the estimates made in the first part of the study would hold for different hues. The total range of physical brightness values used before could not be obtained because of the density of the color filters themselves. The filters were Wratten Nos. 61 (green), 29 (red), and 45 (blue). Again, 10 estimates were made for each brightness value by each *O*.

Observers

Four male college students served as *O*s. All of them had normal visual acuity or had glasses correcting such defects, but one man was judged to be color deficient (deuteranopic) on the basis of the Ishihara test.

Results

Estimates of ½ and ¼ brightness.—The results show in general a remarkable ability to make consistent fractional estimates of brightness, especially at low brightness levels. Even at higher levels these judgments are much more consistent than one might expect. Table I shows the means and standard deviations for the ½ estimates for each of the *O*s and for the combined results of all four. Table II gives the same information for the ¼ judgments.

TABLE I

½ Brightness Estimates

(All values are in millilamberts.)

Observer	Standard Brightness	0.010	0.032	0.10	0.32	5.0	10.0	50.0	100.0
Ge	Mean of Estimates	0.002	0.007	0.021	0.05	0.54	0.86	2.17	3.38
	σ	0.0012	0.0026	0.007	0.012	0.18	0.37	1.15	1.57
Mc	Mean of Estimates	0.004	0.012	0.04	0.10	1.19	1.70	3.92	7.28
	σ	0.0010	0.0020	0.005	0.017	0.27	0.71	0.98	2.83
Ob	Mean of Estimates	0.003	0.008	0.02	0.08	0.70	1.01	2.33	3.87
	σ	0.0012	0.0032	0.007	0.019	0.40	0.54	1.24	1.67
Su	Mean of Estimates	0.004	0.014	0.04	0.10	1.07	1.93	5.20	11.44
	σ	0.0011	0.0023	0.007	0.019	0.24	0.51	1.98	4.17
Group	Mean of Estimates	0.003	0.010	0.03	0.09	0.88	1.38	3.40	6.50
	σ	0.0017	0.0038	0.010	0.037	0.40	0.70	1.87	4.25

TABLE II

¾ BRIGHTNESS ESTIMATES

(All values are in millilamberts.)

Observer	Standard Brightness	0.010	0.032	0.10	0.32	5.0	10.0	50.0	100.0
Ge	Mean of Estimates	0.004	0.011	0.04	0.14	1.42	2.20	6.67	14.05
	σ	0.0009	0.0031	0.010	0.028	0.28	0.57	2.78	5.12
Mc	Mean of Estimates	0.005	0.017	0.05	0.18	2.11	3.39	12.07	21.31
	σ	0.0007	0.0021	0.007	0.024	0.36	0.30	2.26	4.52
Ob	Mean of Estimates	0.004	0.014	0.04	0.14	1.14	1.81	4.74	6.82
	σ	0.0006	0.0025	0.008	0.030	0.33	0.67	1.52	2.55
Su	Mean of Estimates	0.005	0.018	0.06	0.17	1.96	3.64	12.08	26.15
	σ	0.0009	0.0032	0.006	0.028	0.28	0.80	2.64	6.85
Group	Mean of Estimates	0.005	0.015	0.05	0.16	1.66	2.76	8.89	17.08
	σ	0.0010	0.0036	0.012	0.032	0.50	1.03	4.02	8.91

The *O* designated as Ob is the deuteranope. It is quite interesting to note how closely his results parallel those of Ge (who had no color defects) except at the higher brightness values for the ¾ estimates. For this reason the deuteranope's results are combined with those of the other three in the group results at the bottom of the table.

The fact that two *O*s (Mc and Su) make consistently higher estimates than the other two is also worthy of note. The former were obviously using a somewhat different criterion for the fractional estimates than were the latter. There apparently was nothing in the way in which these two (Mc and Su) made their judgments which would account for the similarity of their results and for the differences from the other two. Su took a very long time for most of his judgments and when asked for an expression of confidence stated that he believed he could estimate half brightnesses within an error of ±10 percent. Mc, on the other hand, made very rapid judgments and did not express as high a degree of confidence in his results. He made his estimates so rapidly, in fact, that sometimes *E* suggested that he proceed a little more slowly and adjust the variable up and down several more times. The results, however, show Mc to have been more consistent than Su. The other two *O*s made their estimates at about the same speed and expressed about the same degree of confidence in their results. All four agreed that the ½ estimates were easier than the ¾.

The quotidian variability of individuals was considerable in some cases at certain brightness levels. A table showing this variability has not been included because of space limitations. An analysis of variance showed that this day-to-day variability within *O*s accounted for nearly as much of the total variance at most brightness levels as did the between-*O* variability. There was no consistent trend indi-

cated, the within-*O* variability being in some cases greater and in some cases less than that between *O*s.

The differences between the mean values for each *O*'s and for the combined group's $\frac{1}{2}$ and the $\frac{3}{4}$ estimates were found to be statistically significant at greater than the one percent level of confidence in most cases and at greater than the five percent level in the remaining ones. This indicates that there is a difference between various fractional estimates of a given brightness, and that, at least, one-half can be distinguished from three-quarters. The differences between the means of the successive brightness levels judged were also found to be statistically significant. The very least that can be inferred from this fact is that the brightness intervals used can be counted upon to give consistently different brightness sensations.

Fig. 1 shows the half-judgment function and the three-quarter judgment function as obtained from the combined data. The mean values

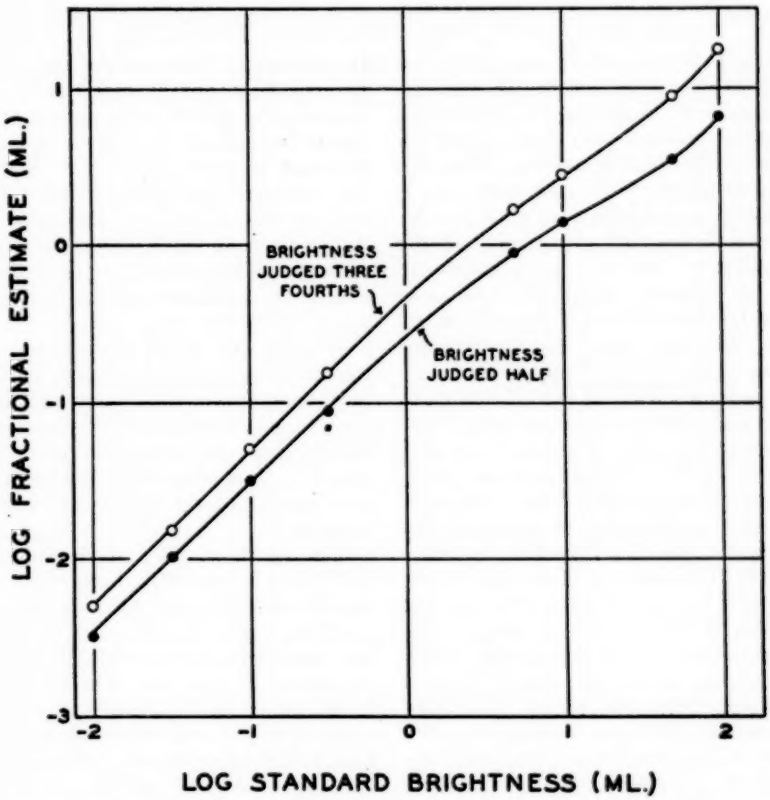

FIG. 1. Fractional estimates of brightness. The filled circles represent the brightnesses which were judged to be half as bright as the standard values, the open circles those judged to be three-fourths as bright.

TABLE III

COMPARISON OF MEANS FOR ½ BRIGHTNESS ESTIMATES BETWEEN FIRST AND SECOND SERIES OF TRIALS

(All values are in millilamberts.)

Observer	Standard Brightness	0.010	0.032	0.10	0.32	5.0	10.0	50.0	100.0
Ge	First Series	0.002	0.007	0.02	0.05	0.54	0.86	2.17	3.38
	Second Series	0.002	0.007	0.02	0.07	1.02	1.12	3.32	5.60
Mc	First Series	0.004	0.012	0.04	0.10	1.19	1.70	3.92	7.28
	Second Series	0.004	0.011	0.04	0.11	1.14	1.81	5.56	7.68
Ob	First Series	0.003	0.008	0.02	0.08	0.70	1.01	2.33	3.87
	Second Series	0.003	0.008	0.02	0.09	1.22	1.71	4.07	4.90
Su	First Series	0.004	0.014	0.04	0.10	1.07	1.93	5.20	11.40
	Second Series	0.004	0.012	0.03	0.13	1.24	1.80	4.68	11.58

from Tables I and II are plotted on logarithmic coordinates for convenience and clearness of exposition. Smooth curves have been fitted by eye to the plotted points. The two curves show a very consistent trend, which can be taken as evidence for the accuracy and consistency of the estimates. There was no indication of any practice effect; in some cases the last day's results showed more, and in some cases less, variability than that evidenced on the first day. As a final check on the reliability of the estimates, four more trials were run 13 days after the end of the first series. Only half-brightness judgments were made this time. No practice on fractionation of any kind had occurred in the 13-day interim. Table III shows the individual means for this last set of estimates, and the means of the first set are included for comparison. The remarkable correspondence of the two sets of data is obvious from the table. The only noticeable difference between the two sets of results is toward making the four *O*s agree more closely with each other. Of course, the second set of data is not very reliable, being based on only four estimates for each *O*, but the similarity between the two is so striking that one is tempted to go beyond the facts and suggest that the estimation of brightness is quite constant for any one person and is not very much different for other individuals. The least that can be drawn from the data is the refutation of any arguments that brightness cannot be estimated at all accurately or consistently.

It may be noted also that the combined raw scores at most of the brightness levels used show a tendency toward normal distributions. There are, however, two cases in which distinct bimodality is exhibited and several of the others are badly skewed.

The effect of flash duration.—As had been expected, there was no significant difference due to the time variable when the duration of both the standard and the variable flashes was reduced to 0.5 sec. On the other hand, some fairly significant differences were found at the lower brightness levels when the duration of the standard was reduced to 30 msec. It is apparent that for such a brief flash of light the sensation of brilliance does not have time to reach its

peak. Troland (6) indicates that for white light this peak is reached in about 0.1 sec. Now, if the sensation accompanying the presentation of the standard does not have time to reach its peak while that for the variable does, we should expect the half-brightness estimates to be lower than they would be under the conditions in the first part of the experiment. In other words, a lower brightness value of the variable would appear to be half as bright, since the effective subjective brightness or brilliance of the standard has been reduced. This is just what the data show. At the four lower brightness levels the mean of each *O*'s estimates is significantly lower for this time pattern than it was for the first part of the experiment. At the higher brightness values, however, the same thing is not true; here one *O*'s estimates were higher for the 30–970 ratio than they were for the 1.3–1.3 ratio in the first part of the experiment. This fact is not surprising, in view of the differential effects in the growth of brilliance due to the stimulus brightness, that is to say, the fact that for a brighter stimulus the sensation should reach its peak in less time than that required for a less bright stimulus. Data obtained by Bills (1) show that this is the case.

Table IV presents the means of the group estimates for the two time patterns used in this part of the experiment and includes also the means of the first set of half-brightness estimates for purposes of comparison. The same data are then shown in graphical form in Fig. 2, where smooth curves have been fitted by eye to the empirical data. No separate curve has been drawn through the points for the 500 msec. duration because these points so nearly coincide with those for the 1300 msec. duration.

It does not seem to go too far beyond the empirical facts to conclude from this part of the experiment that time probably has no significant effect upon the fractionation judgments of brightness so long as the stimuli presented are of sufficiently long duration to allow the sensation of brilliance to reach a relatively steady state.

The effect of hue.—Before anything could be done with the data on hue, some measure of the brightnesses of the colored stimuli was necessary. Although Macbeth readings for colored surfaces are satisfactory for most purposes, the transmission curves for the colored filters, in conjunction with the spectral distribution curve for the phosphor, offered what was judged to be a better basis for more precise determinations. The procedure involved integration of the area under the curve for the phosphor, calculation of the percentage of this

TABLE IV

GROUP MEANS OF ½ BRIGHTNESS ESTIMATES FOR THREE DIFFERENT TIME PATTERNS OF STIMULUS PRESENTATION

(All values are in millilamberts.)

Duration (msec.)		Standard Brightness						
Standard	Variable	0.010	0.032	0.10	0.32	5.0	10.0	100.0
1300	1300	0.003	0.010	0.03	0.09	0.88	1.38	6.50
500	500	0.004		0.03			1.23	6.67
30	970	0.001	0.004	0.01	0.04	0.66	1.09	5.73

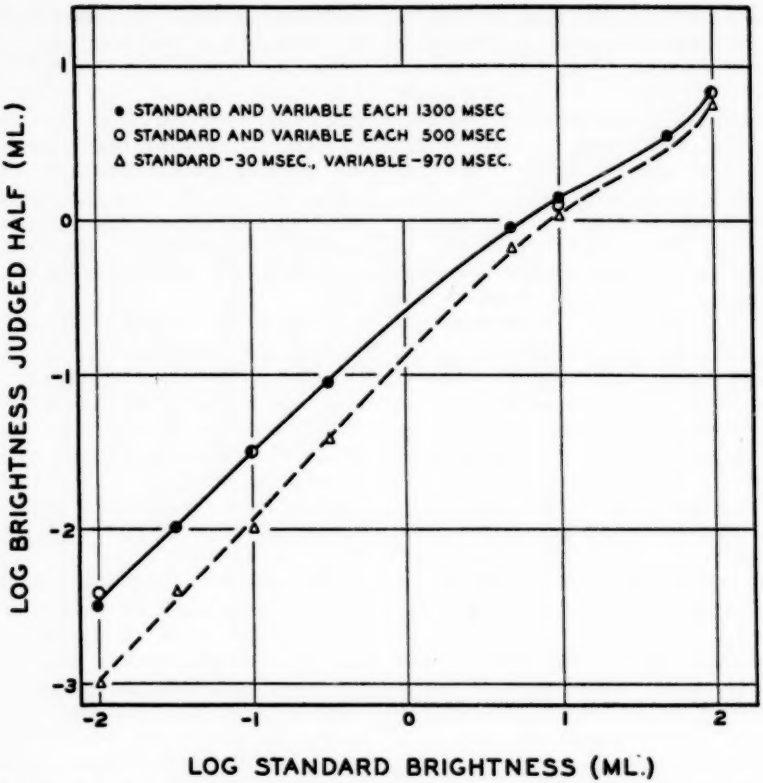

FIG. 2. The effect of duration on the half brightness estimates. The ordinate represents the brightness of the variable which was judged half as bright as the standard (abscissa).

area under each of the filter curves, and translation of these percentages into brightness units. By this method the maximal standard brightness through the red filter (Wratten No. 29) was found to be 8.3 millilamberts, for the green (Wratten No. 61) 15 millilamberts, and for the blue (Wratten No. 45) 8.6 millilamberts.

If any of the calculated brightness values for the filters are incorrect, the worst that can happen is a displacement of the half-judgment function. This is not a particularly serious fault, since the shape of the function would not be changed. If we assume that the curve obtained from the first set of data is valid, then it appears from the results that just such a shift has occurred to a slight degree. Table V and Fig. 3 present these results. On the graph the half-judgment function obtained in the first part of the experiment is shown for purposes of comparison. The mean values through which this curve was drawn are not included in the table.

If we were to displace the curves for the various hues slightly, the upper portions of each curve would very nearly coincide with the curve for the white light. The lower por-

tions, however, depart somewhat from the curve for white, the red curve bending slightly upward away from the white while the curves for green and blue turn generally downward. The Purkinje shift is apparently showing up in these results. One would expect on the basis of this phenomenon that as lower brightness values at the short-wave end of the spectrum were approached, less and less physical brightness would be necessary to make any subjective estimate seem like half of a higher brightness value. The reverse would, of course, be true of the values at the long-wave end of the spectrum.

It is of extreme interest to note the slight effect of hue evidenced in the results for the color deficient individual. One might expect that his estimates would vary considerably from those for white light, especially when green was used. While there is some evidence of a shift in his results toward closer conformity with the results of the other *O*s, i.e., a tendency for his estimates to be slightly higher than formerly, the effect is negligible and is not at all consistent from one brightness level to another. The explanation for this would seem to be in the nature of the problem, which is a judgment of relative values for the same hue. If the deuteranope's sensitivity to green, for instance, is reduced, this might serve only to shift his judgments slightly on the curve, since both the standard and the variable would be affected to nearly the same relative degree.

Although no statistical tests for the significance of the differences between the estimates for hue and those for white light could be readily applied, since the standard brightness values were not the same, the differences obtained may be significant, the Purkinje phenomenon being such a well-established fact. How, then, will this evidence alter our conclusions about the generality of fractional estimates of brightness? The obvious implication is that we should limit our conclusions either to white or a constant hue, or if we wish to compare white light with colored, we should probably exclude the range of brightnesses in which the Purkinje phenomenon shows up significantly.

TABLE V

GROUP MEANS FOR ½ BRIGHTNESS ESTIMATES FOR VARIOUS HUES

(All values are in millilamberts.)

Wratten No. 29 (Red)								
Standard Brightness	0.008	0.026	0.415	0.83	4.15	8.30		
Mean of Estimates	0.004	0.011	0.148	0.29	0.99	1.50		
Wratten No. 61 (Green)								
Standard Brightness	0.002	0.005	0.015	0.048	0.75	1.50	7.50	15.0
Mean of Estimates	0.0004	0.002	0.004	0.014	0.25	0.44	1.38	1.84
Wratten No. 45 (Blue)								
Standard Brightness	0.0009	0.0028	0.0086	0.028	0.43	0.86	4.30	8.60
Mean of Estimates	0.0003	0.0009	0.0023	0.007	0.12	0.24	0.99	1.53

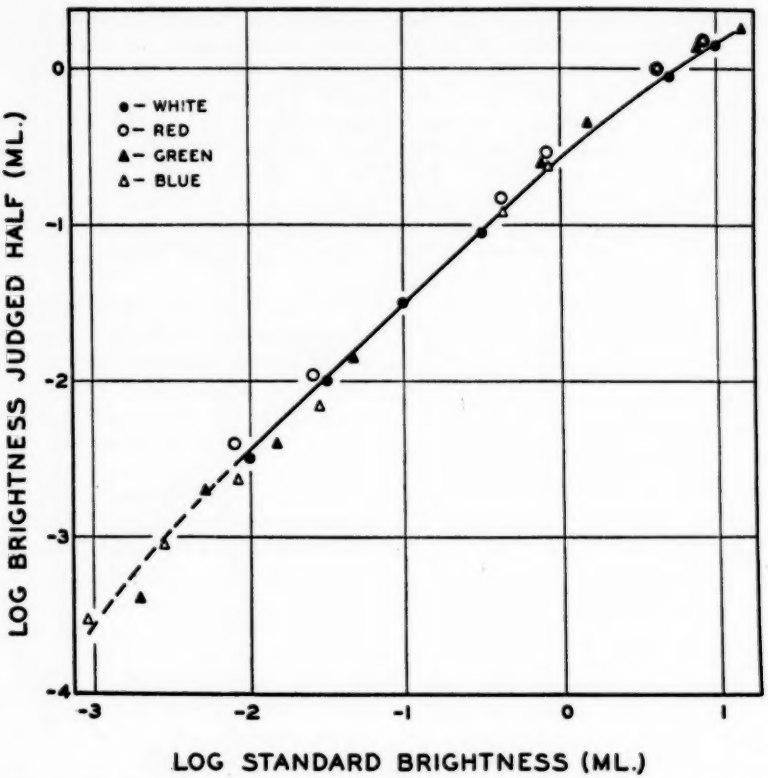

FIG. 3. The effect of hue on the half brightness estimates. The ordinate represents the brightness of the variable which was judged half as bright as the standard (abscissa). The solid line is the half-judgment function from Fig. 1, shown for purposes of comparison. The dotted portion of the curve is an extrapolation based on the estimates for the various hues.

The Brilliance Scale

From the half-judgment function in Fig. 1 a brilliance scale was constructed after the manner used by Stevens (5) and described again by Reese (3). The procedure was briefly as follows. One hundred subjective brightness units were arbitrarily chosen as equal to one millilambert. This point was plotted first in Fig. 4. Then from Fig. 1 it was found that 0.26 millilamberts appeared to be half as bright as one millilambert; therefore, 0.26 millilamberts was assigned the value 50 or $\frac{1}{2}$ of 100 brilliance units. In a similar manner the physical value to be assigned to the brilliance number 25 was determined, etc. The value from Fig. 1 which appeared twice as bright as one millilambert was six millilamberts; this value was, therefore, assigned the brilliance number 200.

A brilliance scale constructed in the same manner from the $\frac{3}{4}$-judgment function does not agree as closely with the scale just constructed as one might wish. With one millilambert

set equal to 100 subjective brightness units again, the points fall slightly above the lower part of the scale in Fig. 4 and below the upper part, the discrepancies becoming progressively greater at the higher brightnesses. In other words, it is somewhat as though the curve were rotated slightly clockwise about its anchor point. Upon questioning, the *O*s reported that they had been making the ¾ estimates in the following manner. Picturing first what they believed to be half of the interval to be fractionated, they then estimated ½ of the difference between this imaged midpoint and the standard. A different type of judgment would therefore seem to be involved in this case, with additional errors due to the estimation of the imaged ½ brightness. In making the ½ estimates, *O* has two real brightnesses and one imagined one, the zero point, to work with. The ¾ judgment apparently has in addition a second imagined brightness, at least for the way in which this estimate was made by the present

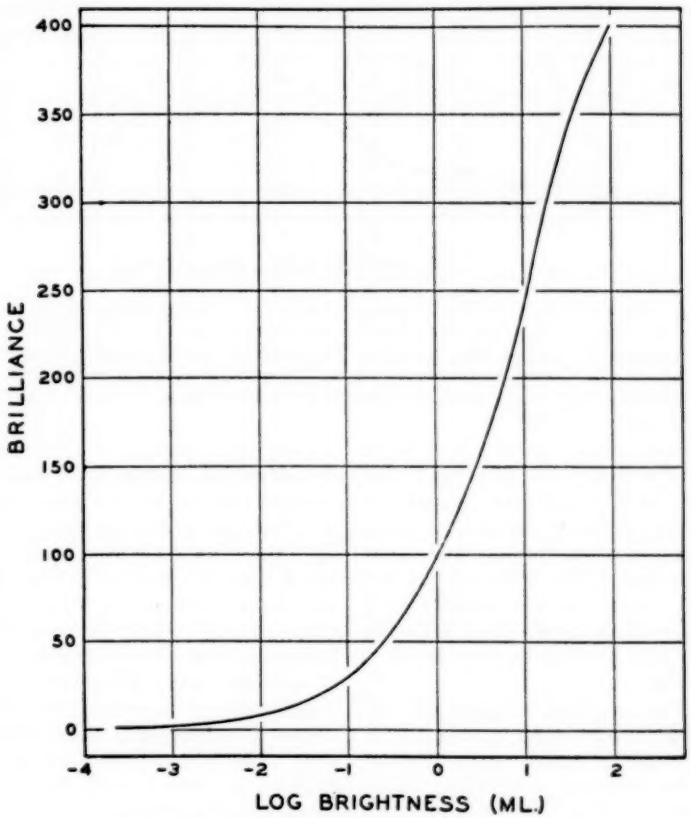

FIG. 4. The brilliance function. Subjective brightness or brilliance is plotted as a function of physical brightness.

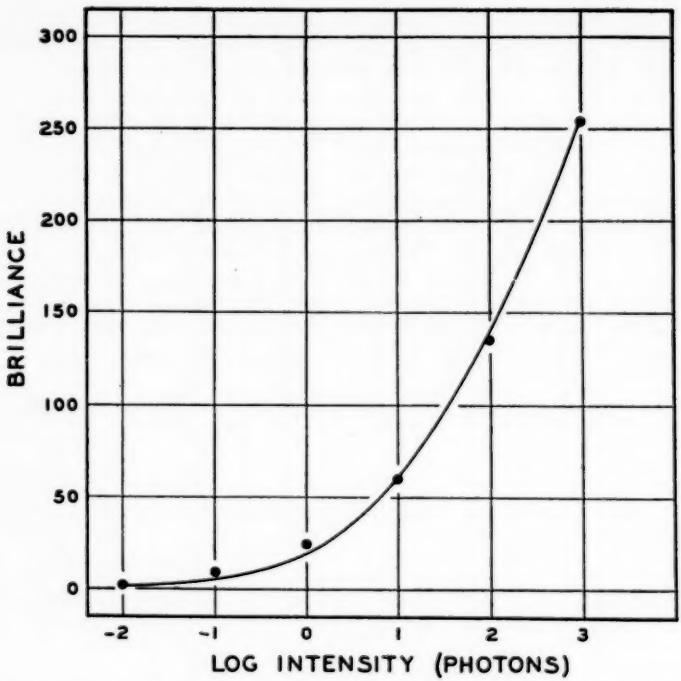

FIG. 5. The relation of the brilliance function (solid curve) to the integrated DL's (solid circles) for brilliance. The numerals along the ordinate represent the number of DL's as a function of the intensity of retinal illumination (after Troland). The conversion factor for changing from these DL units to brils is 1.61. Since the points representing Troland's values were scaled from the graph published in the reference mentioned, the accuracy of their location is, of course, subject to the limitations of such a method.

*O*s. It is quite possible that this imaged half brightness would be somewhat different from the estimates made when the half judgment was required. The influence of the variable when being adjusted for $\frac{3}{4}$ could have had the effect of lowering the imagined $\frac{1}{2}$ and consequently, also, the $\frac{3}{4}$ judgment.

Another possible explanation for the differences between the two scales lies in 'naming' errors in connection with the required fractions. Both these possibilities are substantiated by the fact that if the $\frac{3}{4}$ judgments are treated as though they had been $\frac{2}{3}$ judgments, the resulting points fall almost exactly on the curve in Fig. 4. For the above reasons, plus the fact that the scale constructed from the $\frac{1}{2}$ judgments agrees so closely with the integrated DL's for brilliance (see Fig. 5), the scale in Fig. 4 is taken as being a better representation of the *O*s' experiences than the $\frac{3}{4}$ judgment would be. The discrepancies do, however, raise very important questions from the standpoint of scaling theory and need to be investigated thoroughly.

As stated above, the $\frac{1}{2}$ judgment scale agrees very closely with the

brilliance function obtained by Troland (6) from the integration of DL's. The relationship of the two sets of data is shown in Fig. 5. In this graph photons are represented on the abscissa and integrated DL's on the ordinate. The brilliance function of the present experiment is shown on these axes as the solid curve. To plot this curve on Troland's coordinates it was necessary to convert our units into DL units. In comparing the two, it was found that a simple constant (1.61) multiplying the DL units at any selected brightness would give satisfactory agreement with the number of units assigned by the present method to that same physical brightness. In other words, if our units were plotted against integrated DL's, the plotted points could be fitted satisfactorily with a straight line having a slope of 1.61. This means that the two functions are essentially the same but that the units are smaller on one scale. The close fit of Troland's points (filled circles) and the curve in Fig. 5 indicates that, provided our function is valid, the DL's for subjective brilliance are all approximately equal up to 100 millilamberts of physical brightness and may be validly integrated to obtain measures of brilliance.

It is unfortunate that the limitations of the apparatus prevented the exploration of subjective estimates of brightnesses above 100 millilamberts. The ease of presentation and control of the stimulus with this kind of equipment gives it an advantage over many types of optical arrangements. The work will obviously not be complete until the higher brightness values are investigated.

The term bril is tenatively suggested as a name for the unit of brilliance. Wright (8) has constructed an arbitrary subjective brightness function to correspond to the way in which the difference limen for brightness varies with intensity.[3] The unit of this scale he has named the 'brill.' The same name had been thought of for the present scale unit, but in order to differentiate the two it seems better at the present time to shorten the word to 'bril' for our scale.

A possible criticism of the experiment lies in the fact that the method of presentation employed was one of successive rather than simultaneous comparison. Such a method allows the possibility of a time error. If such an error has been introduced, it would, of course, bias the obtained fractionation estimates somewhat, i.e., a positive time error would tend to make these estimates higher and a negative one lower. However, in view of the nature of the total pattern of stimulation, it seems quite unlikely that any time errors have seriously affected the results. One might even conceive of a positive and a negative time error both operating to some extent and cancelling each other.

[3] The following is Wright's description of the manner of constructing his scale: ". . . while $\Delta B/B$ is approximately constant at medium and high values of B, at low brightnesses $\Delta B/B$ increases as B decreases. If, then we consider ΔB as an elementary contrast step, we can see that a logarithmic scale of photometric brightnesses would provide a satisfactory representation of subjective brightness values at high values of B, but at low values the scale would be too open. Conversely, a linear scale of photometric brightness might be reasonable at low brightnesses but would be too open at high values.

"As a compromise, the author [Wright] proposed the following scale and definition:

"When an extended surface of photometric brightness B is viewed by a dark-adapted observer using both his eyes, then the apparent or subjective brightness B_s should be calculated from the equation

$$B_s = 10 \log_{10} (B + 1).$$

The unit of apparent brightness may be called a 'brill,' and B_s will be expressed in brills when the unit of B is 10^{-3} equivalent foot-candles."

In other words, the brief exposure to the higher brightness of the standard might result in slight adaptation effects giving rise to a positive time error which could be cancelled by the negative time error interjected due to the disparity in time between the presentation of the standard and that of the variable.

Summary

The fractionation technique has been used to obtain estimates of brightness for various time patterns and for various hues. Because of equipment limitations only the range from about 0.0001 to 100 millilamberts has been investigated thus far.

It has been shown that such estimates can be made consistently with only a reasonable degree of error.

The duration of the stimulus has been found to have no appreciable effect on these judgments so long as the stimulus duration is long enough to allow the sensation to reach a steady state.

Hue, likewise, has been shown to cause little change except, perhaps, at very low levels.

A subjective brightness or brilliance scale has been constructed and has been found to agree closely with the curve obtained by integrating DL's.

(Manuscript received
September 17, 1948)

References

1. Bills, M. A. The lag of visual sensation in its relation to wave lengths and intensity of light. *Psychol. Monogr.*, 1920, **28**, No. 5.
2. Hanes, R. M., & Williams, S. B. Visibility on cathode-ray tube screens: The effects of light adaptation. *J. opt. Soc. Amer.*, 1948, **38**, 363–377.
3. Reese, T. W. The application of the theory of physical measurement to the measurement of psychological magnitudes, with three experimental examples. *Psychol. Monogr.*, 1943, **55**, No. 251.
4. Stevens, S. S. A scale for the measurement of a psychological magnitude: loudness. *Psychol. Rev.*, 1936, **43**, 405–416.
5. Stevens, S. S., Volkmann, J., & Newman, E. B. A scale for the measurement of the psychological magnitude pitch. *J. acous. Soc. Amer.*, 1937, **8**, 185–190.
6. Troland, L. *Psychophysiology. II. Sensation.* New York: D. Van Nostrand, 1929.
7. Williams, S. B., Bartlett, N. R., & King, E. Visibility on cathode-ray tube screens: Screen brightness. *J. Psychol.*, 1948, **25**, 455–466.
8. Wright, W. D. *Researches on normal and defective color vision.* London: Henry Kimpton, 1946, Ch. XXIII.

VISUAL RESOLUTION WITH PERIODICALLY INTERRUPTED LIGHT

BY VIRGINIA L. SENDERS

Wellesley College *

Introduction

Visual acuity has often been called 'a form of brightness discrimination' (2, p. 34) or of intensity discrimination (8, 13). Such statements usually imply, not only that there must be some discriminable difference between the dark and light portions of a pattern in order for that pattern to be resolved, but also that the fundamental physiological processes underlying visual acuity and brightness discrimination are the same. The fact that visual acuity and brightness discrimination have been shown to vary in a similar way with a variation in intensity has been adduced, particularly by Hecht and his collaborators (6, 7, 8, 9, 10, 13), as additional evidence for this point of view.

However, if any conditions can be found where visual resolution and brightness discrimination are different functions of the same independent variable, it is then no longer permissible to consider one to be a different aspect of the other. Visual acuity, while it may be related to brightness discrimination, could then no longer be considered to be merely an aspect of that process. Recent experiments by Crozier and Wolf (5) indicate that such divergent functions can be obtained by using interrupted rather than steady light, and measuring the threshold of resolution as a function of light-time fraction in the interruption cycle. These experimenters were seeking to determine the effect of subdividing the test field on the form of the flicker contour. However, in the course of obtaining the flicker contours, reports were also made on the subjective appearance of the test field. It was noticed that with certain kinds of subdivision, the intensity (whether fused or producing flicker) required for a pattern to be recognized was considerably less than would have been predicted on the basis of the Talbot-Plateau Law. This law states that when the light is interrupted, the fused brightness is the same as might be exptected if the same light energy per cycle were equally distributed throughout the entire flash cycle. That is: if the flash-intensity is kept constant, the decrease in fused brightness is proportional to the decrease in the percent of light-time in the cycle (12, 14). Crozier and Wolf found that, in some cases, the flash-intensity required for resolution remained constant while the light-time fraction was decreased to as little as one-tenth. Thus it appeared that a situation had been found in which the energy might be decreased to as little as one-tenth of the original level (with a corresponding decrease in brightness) without producing any change in resolution. It was suggested by these authors that the use of interrupted light might provide a technique for the separation

* This study was performed as part of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology, Harvard University. The writer is indebted to Professor William J. Crozier of the Laboratory of General Physiology for his direction of the investigation; the responsibility for the manuscript in its final form, however, rests entirely with her.

of effects due to *intensity* (a dimension of the physical stimulus) from those due to *brightness* (an attribute of the sensation).

The Talbot-Plateau Law is a statement about brightness (i.e., not about brightness discrimination). But if visual acuity is a form of brightness discrimination it seems reasonable to assume that the greater the brightness differential between the illuminated and nonilluminated portions of the image, the better will be the spatial discrimination. In other words, if light bars on a dark field are to be discriminated, it would seem reasonable to assume that they should always be equally bright when they are resolved. A brightness discrimination of some sort is certainly being made, since until some brightness difference between portions of the field appears, the observer certainly cannot resolve any bars. But the question is: will the same brightness (or the same brightness difference between the illuminated and nonilluminated portions of the field) always be required for the observer to reach the same criterion? Will the intensity of light required to produce resolution of a pattern be the same as that which would be required to produce a brightness equal to that of the steady light?

The purpose of the present investigation was to make a preliminary study of the dependence of the resolution threshold on light-time fraction when the light is periodically interrupted, and to compare the nature of the function with the brightness function as expressed in the Talbot-Plateau Law.

Apparatus

The apparatus used was the Visual Discriminometer, which has previously been described (3, 4). Essentially the Discriminometer consists of a tungsten lamp with a flat ribbon filament. The lamp is operated on 28 amperes of alternating current at six volts. It provides a constant source of illumination brought by collimating lenses to a focus where it may be interrupted by motor-driven sector-disks; the light beam pases through slits where a target may be inserted and is finally focussed by the *O* by means of a microscrope headpiece. Another beam from the same source passes through a tiny hole, and the image of the small illuminated hole provides a fixation point. Illumination intensity is controlled by the use of neutral filters and by a wedge and balancer. Wavelength composition is altered by the use of colored filters. The light-time fraction is controlled by the use of a series of sector-discs with suitably disposed openings. It shoud be noted that the light-time fraction is here defined as the *percent of light in the total cycle*. It is $t_L/(t_L + t_D)$ and is not to be confused with the ratio between the light-time and the dark-time, t_L/t_D.

The target is a specially prepared photographic film (cf. 5) inserted in the slits of the Discriminometer. Accurate focussing for any wave-length is assured by the *O*'s control of the microscope headpiece.

The fixation point is red and of such spectral composition and intensity that it appears as red only in the fovea.

Procedure

The procedure in each experimental session was as follows:

1. The *O* was dark-adapted for 30 min. preceding experimentation.

2. The *O* fixated the red fixation point and signalled to *E* that he was ready to begin the experiment. At this time, the *O* actually saw a dark field with a dimly illuminated square in it.

3. *E*, by movement of the optical wedge, increased the intensity of the test field until *O* signalled that he could barely distinguish the bars of the target. To the *O*, the square field appeared to become increasingly brighter, and at some point striations became visible. When the flash frequency was low enough, the test patch appeared to the *O* to flicker.

The exact nature of the resolution endpoint varied from *O* to *O*. LHC, for example, did not signal until the bars had become quite clear and distinct; DLP required only that some evidence of striation be barely discernable, and NHP did not signal until he could tell with certainty that the bars were vertical. The *O*s were requested to keep their respective criteria as constant as possible from one session to the next, and it is

believed that they were able to do this fairly successfully. However, when the observations in the periphery were started, LHC was forced to relax her criterion slightly, since the clarity which she had required in the fovea could not be obtained in the periphery.

4. When the *O* signalled by pressing a telegraph key that his criterion had been reached, *E* kept the intensity at the same level for a few moments in order that the *O* might become adapted to that intensity level. He then lowered the intensity below the resolution threshold, gave the *O* a 'ready' signal, and increased the intensity until the *O* again signalled. After taking a few exploratory readings, recording was commenced and readings were taken until ten consistent readings had been obtained.

5. When 10 readings had been obtained, the light-time fraction was changed by changing the sector-disc. Light-time fractions used were: steady light (1.00), (occasionally 0.90), 0.75, 0.60, 0.50, 0.25, 0.10, and 0.03. Ten consistent readings were taken at each of these light-time fractions.

6. Two such series were obtained under each set of experimental conditions. One was a descending series, going from steady light to 0.03; the other was an ascending series going from 0.03 to steady light. This was done in order to take account of any possible effect of fatigue or adaptation. Where possible, the two series were run in immediate succession, the *O* dark-adapting for 15 min. between series.

Three *O*s were used extensively in the experiment.[1] LHC was a research assistant in physiology. She was 40 years old and had emmetropic eyes and normal color vision. She was able to observe under more regular conditions than either of the other two *O*s. NHP was a graduate student in biology; he was aged 27 years and had pronounced exophoria, slight astigmatism, and normal color vision. DLP was a graduate student in psychology. She was aged 23 years and had myopic eyes and normal color vision. The fact that all the *O*s were not emmetropic was not considered a disadvantage, since corrections could be made by adjusting the eyepieces of the Discriminometer.

A few sample series were obtained on each of four other *O*s. The results of these experiments are entirely consistent with the results of the three principal *O*s, but since there are not enough experiments to make possible the comparison of results under different conditions, these data are not presented.

[1] The writer is greatly indebted to Mrs. Louise H. Crozier, Mrs. Dorothy L. Postman, and Nathan H. Pulling, who served as *O*s in these experiments.

Parameters

1. Flash frequency. The flash frequencies tested were 4, 6, 8, and 10 flashes per sec., slightly above fusion (about 25–30 flashes per sec.) and very much above fusion (89 flashes per sec.).

2. Size of test patch. The test patch used throughout the experiment was a grid of alternate light and dark bars, each bar subtending 6.6′ of arc at the retina. Test patches were 2° square (nine dark bars), 1° square (four dark bars), and 1° × 2° (four dark bars 2° in height). It was also possible, by using a different ocular, to decrease the width of the bars, along with the size of the test patch, by approximately one-half. This was done in a few cases.

3. Retinal location. The target was presented either in the fovea or displaced 10° temporally from the fovea in the case of LHC and DLP, or displaced 10° nasally in the case of NHP.

4. Wave-length composition. White light, red light, and blue light were used. For the red light, Wratten filter number 29 was used, and for the blue light, Wratten filter number 47.

It was not possible to obtain observations with all possible combinations of these parameters, but enough different combinations were used to get a good idea of the shape of the function under each condition.

It should be noted that, whatever the parameters, the dependent variable was always the same: intensity required for resolution. The visual angle was *not* changed. Since visual acuity may depend upon image area, it is particularly desirable that this should be kept constant in comparative series. This requirement has not usually been observed by earlier experimenters.

Results

For each of the single series of 10 observations, the median intensity in millilamberts was calculated. The median of the combined distribution, which included both the ascending and the descending series, was also calculated.[2] The median rather than the mean was used because it repre-

[2] For complete tabular summaries of the data order Document 2649 from American Documentation Institute, 1719 N Street, N.W., Washington 6, D.C., remitting $0.50 for microfilm (images one in. high on standard 35 mm. motion picture film) or $0.70 for photocopies (6 × 8 in.) readable without optical aid.

sents the 50 percent point of a psychometric function. The median of a series of logarithms is, of course, the log geometric mean of the series, if the series is distributed normally.

A sample function is presented in Fig. 1. Here the logarithm of the intensity (millilamberts) required for resolution has been plotted against the logarithm of the light-time fraction. Each point is the median of 20 observations. The reason that a double logarithmic plot has been used is to facilitate the comparison of the intensity required for resolution with the intensity which should, according to Talbot's Law, be required to make a brightness match with the steady light. In this figure a 45° line, which is the contour of constant brightness as predicted by Talbot's Law, has been drawn. When the points representing the intensity required for resolution lie significantly below this line, we must conclude that in some way the interrupted light is more efficient in producing resolution than it is in producing brightness. Of course, when the flash frequency is below the critical fusion frequency, the brightness is not described by the Talbot-Plateau Law. However, the contour has nevertheless been presented for comparative purposes. A consideration of the effect of a possible brightness enhancement at

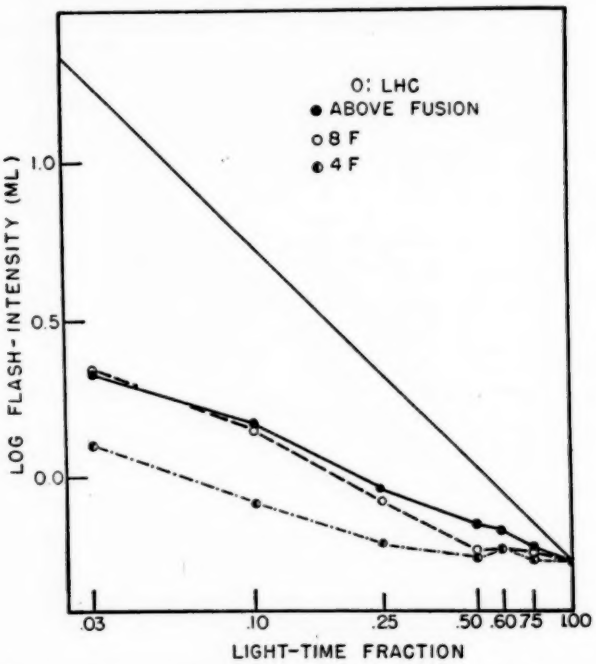

FIG. 1. The relationship between the logarithm of the flash-intensity required for resolution and the logarithm of the light-time fraction, for different flash frequencies. A 2° × 2° test patch, centered at the fovea, is used with red light. All data are expressed as multiples of the steady light threshold for the series obtained with a flash frequency above fusion. The 45° line is the contour for the constant Talbot (fused) brightness.

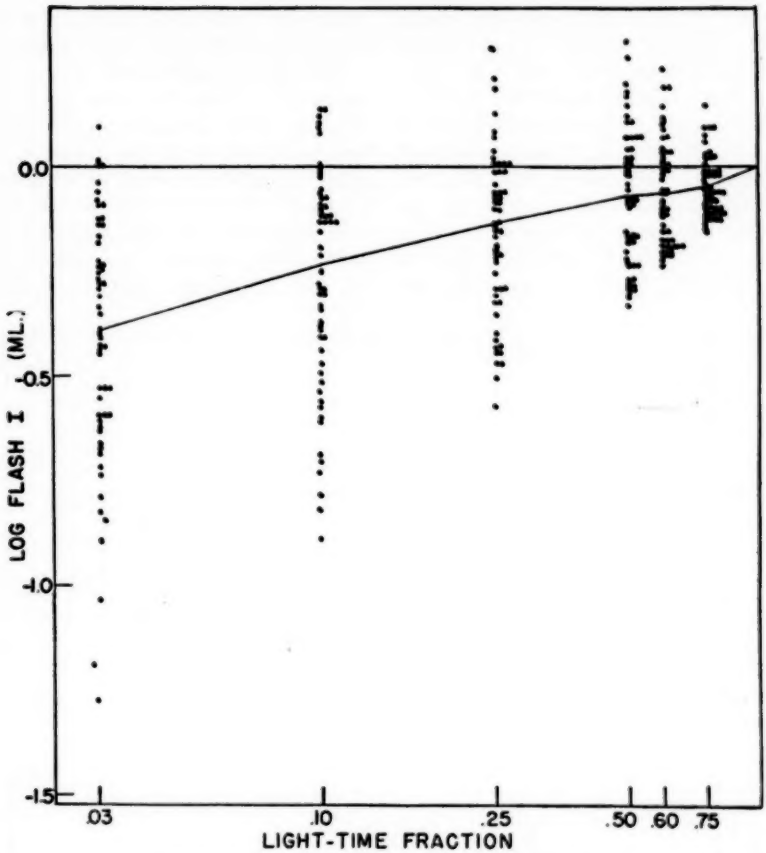

FIG. 2. Scatter diagram showing deviations from the Talbot brightness contour as a function of the light-time fraction. The results are for three observers, when the flash frequency was above the critical fusion frequency. A line has been drawn to connect the median deviation at each light-time fraction. The horizontal line represents the contour for the constant Talbot (fused) brightness.

low flash frequencies on the intensity required for resolution is deferred to a later point.

The data are summarized graphically in Figs. 2, 3, and 4. In these figures, scatter diagrams have been presented showing, for each combined series of 20 observations, the deviation of the intensity (millilamberts) required for resolution from the intensity predicted on the basis of Talbot's Law.

From these data it may be seen that, in general, the intensity required for resolution is less than would be required, on the basis of the Talbot-Plateau Law, to make a brightness match with the steady light. There are some exceptions to this statement, but an examination of Figs. 2, 3, and 4 will show that there need be no question of the significance of the deviations. In some way, interrupted light is more effective in producing

resolution than in producing brightness. Brightness and resolution are different functions of the same independent variable, light-time fraction. (Incidentally, the subjective reports of the *O*s confirm these statements. When the light-time fraction was decreased and less energy was required for resolution than would have been predicted by Talbot's Law, one *O* remarked: "It seems too dim, but somehow I can see the stripes clearly." Such comments were fairly frequent.)

Individual differences were pronounced. Under the same conditions there is little similarity between the resolution contour for different *O*s, either in absolute level or in general shape. This fact should be neither as surprising nor as discouraging as it may at first appear. The exact form of the resolution contour depends

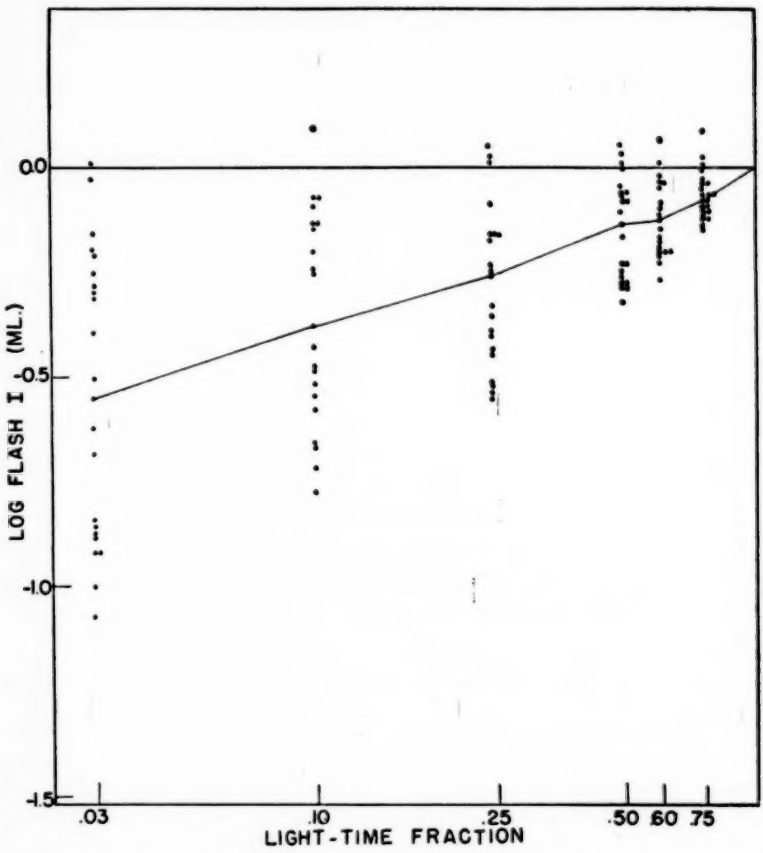

FIG. 3. Scatter diagram showing deviations from the Talbot brightness contour as a function of the light-time fraction. The results are for three observers, when the flash frequency was 8F/sec. A line has been drawn to connect the median deviation at each light-time fraction. The horizontal line represents the contour for the constant Talbot (fused) brightness.

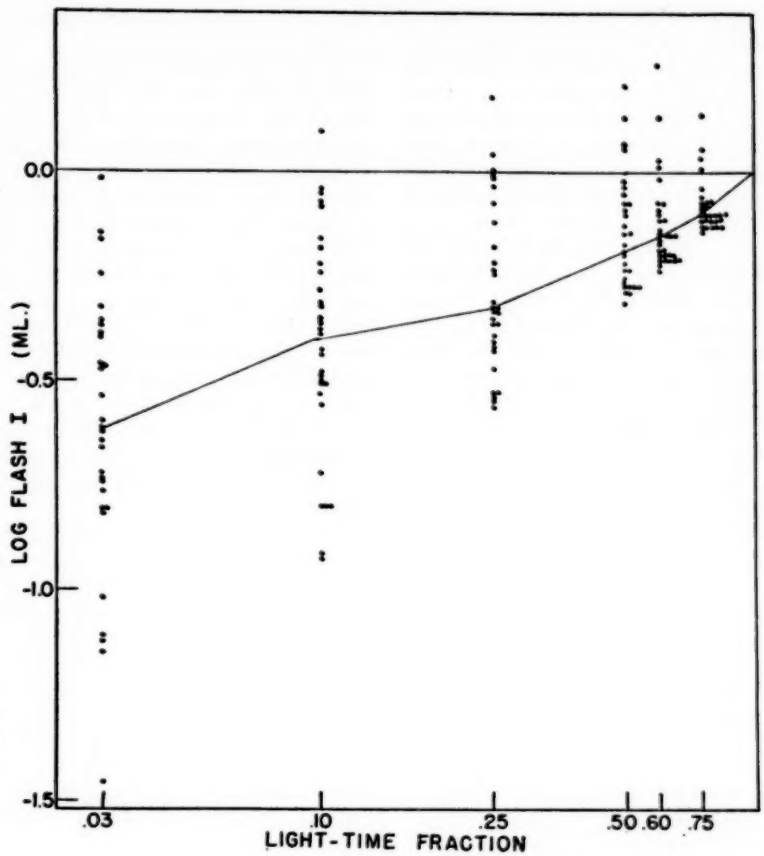

FIG. 4. Scatter diagram showing deviations from the Talbot brightness contour as a function of the light-time fraction. The results are for three observers, when the flash frequency was 4F/sec. A line has been drawn to connect the median deviation at each light-time fraction. The horizontal line represents the contour for the constant Talbot (fused) brightness.

upon many variables, only a few of which are controlled in the laboratory. One such variable, which may be very important indeed, is the natural periodicity of the body. Pronounced individual differences in eye-movements, neural recovery cycles, possibly the natural rhythm of the cortex, or the retinal pulse, may exist, and may have influenced the results, but it was not possible to investigate these factors at this time. Regular and marked individual differences in the form of the flicker contour have been reported by Crozier and Wolf (4) and the possible effect of an age difference was noted by them.

Certain general tendencies in respect of the four parameters used in this experiment were noted. These will be discussed in turn.

The data at hand show no regular

relationship between flash frequency and the form of the resolution contour. However, an examination of Figs. 2, 3, and 4 show that there is a general tendency for greater deviations from Talbot's Law to occur when the flash frequency is low than when it is high. For comparative purposes the results may be divided into two groups: those obtained when the flash frequency was above fusion, and those obtained when it was below fusion. In the first case, as has been stated, the brightness is adequately described by the Talbot brightness contour. But this contour is not applicable in the second case because of the Brücke-Bartley effect; in the case of flash frequencies below fusion we should expect that lower flash-intensities would be required to match the brightness of the steady light than those predicted by Talbot's Law. Thus, if the flash-intensity required for resolution depended only upon brightness, as a brightness-discrimination theory of visual acuity would suggest, th.n the resolution contours obtained with flash-frequencies below fusion should lie below those obtained when the light is subjectively fused.

It may be seen at once that the lines and points for the data obtained with flash frequencies of four and eight flashes per sec. tend to lie below those obtained with a flash frequency above fusion. This is what we should expect if visual acuity improved in a regular way with an increase in brightness. But in this case we should also expect that, since the Brücke-Bartley effect [according to Bartley (1)] is maximum at eight to twelve flashes per sec., the deviations from the Talbot contour would also be greatest at this flash frequency. This is not the case. The median deviations are greatest with a flash frequency of *four* flashes per sec., next for eight, and least for flash frequencies above fusion.

This finding is entirely in accord with certain unpublished data obtained by Crozier and Wolf. These experimenters determined the flash-intensity required for resolution of a pattern as a function of flash frequency. Their results show that the flash-intensity required for resolution shows a positively accelerated increase in flash-frequency up to a certain point, beyond which it remains constant. No discontinuity in the curve was apparent at flash frequencies of eight to twelve flashes per sec., as would be expected if the enhanced brightness due to the Brücke-Bartley effect produced an enhancement of resolution.

When the flash frequency is very high (89 flashes per sec.) the flash-intensity required for resolution is either approximately the same as that required to produce a brightness subjectively equal to that of the steady light, or else it is slightly higher. In other words, the resolution contour lies either very close to or slightly above the brightness contour. The points, however, do not lie on a straight line, as do those of the brightness contour.

When we compare the magnitude of the deviations of the resolution contour from the brightness contour for different sizes of test patch, we find that when the test patch is centered in the fovea there is some tendency for the contour for the smallest field to have the largest negative deviations. It must be remembered that we are comparing not only image areas, but also fields which have a different number of stripes, and fields which have the same number of stripes but in which the stripes are of different length. The perimeter of light-dark separa-

tion is also changed. When the test patch is centered in the periphery of the retina the contour for the 1° × 1° field also lies farthest below the brightness contour in almost every instance where a consistent difference is evident, and the resolution contour for the 1° × 2° field shows a surprising tendency to lie close to the brightness contour or even to cross it or to lie above it. Quite consistently the contour for the 2° × 2° field lies between that of the 1° × 1° field and that of the 1° × 2° field.

The influence of wave-length composition on the deviation of the resolution contour from the brightness contour depends upon the retinal location of the test patch. For test patches centered in the fovea, contours obtained with red light lie below those obtained with blue or white light. That is, red light produces a greater difference between the brightness and resolution contours than does blue light or white light. This finding held, in general, for all *O*s, although not for all observations.

When the test patch is centered in the periphery of the retina, the negative deviations of the resolution contour from the Talbot brightness contour are smallest for red light. The resolution contours for blue and for white light lie below those for red light in the majority of cases. The difference between the resolution contours for different wave-lengths is not as marked in the periphery as it is in the fovea.

In most cases the magnitude of the negative deviations of the resolution contour from the brightness contour is larger when the test patch is centered in the periphery than when it is centered in the fovea. The exact magnitude of the differences depends upon image size, flash frequency, and wave-length composition. With white light, there is a large and consistent difference between the resolution contours in the fovea and in the periphery. This difference is always in the same direction and appears in every case. Always the resolution contour for test patches centered in the periphery lies far below the contour for test patches centered in the fovea. The effect is much the same when blue light is used instead of white. In general, the greatest difference appears with the smallest fields. With red light, there is no such consistent difference between the fovea and the periphery. The deviations of the resolution contour from the brightness contour are usually approximately equal, and what differences do appear are not in a consistent direction. Often the shape of the resolution contours is also similar.

Discussion

The results of the present investigation indicate that a situation has been found in which visual acuity and brightness discrimination can be shown to be different functions of the same independent variable. Therefore, they cannot be identical processes. This is the old argument that there may be more sensory attributes than stimulus dimensions. Just as loudness and volume are two different sensory attributes which are both dependent upon the stimulus dimensions of intensity and frequency in the auditory stimulus, so brightness discrimination and visual acuity may be different 'attributes' of the visual sensation and may depend in different ways upon the spatial and temporal distributions of intensities in the visual stimulus. It is not necessary to demonstrate a different physiological correlate for each 'attribute' in order to show that the discrimination made by the *O* exhibits a different depend-

ence upon the physical dimensions of the stimulus.

However, it is interesting to speculate upon the physiological mechanism which enhances resolution without enhancing brightness when the light is interrupted. A mechanism which might account for the above findings will be suggested, but it remains at this point only a speculative possibility, requiring experimental refutation or verification.

There are two main groups of factors which might influence visual resolution: (1) those factors which influence the formation of the retinal image—the actual distribution of intensities on the retina; and (2) those factors within the visual projection system, from the retina to the cortex, which influence the final pattern of impulses arriving at the visual cortex. Within the first category are included diffraction by the pupil, the scattering and diffraction of light by the ocular media, chromatic and spherical aberration, and physiological nystagmus. Of these, we may safely dismiss the first three; there is no way in which a periodic interruption of the light could change its diffraction, scattering, or aberration. The effects of physiological nystagmus, a random, saw-tooth movement of the eyeball, may be profound, as will be suggested, but not at a retinal level. Since it is a fundamental principle of probability theory that a place selection does not alter the relative frequencies of events in a collective, we may conclude that the relative distributions of intensities on the retina for steady light and for interrupted light, when the exposure is of a reasonable duration, will be the same. In general, we may therefore conclude that interrupting the light does not alter in any way the relative distribution of intensities on the retina. Whatever factors operate to make resolution more efficient with interrupted light than with steady light must certainly be factors operating after the formation of the intensity distribution upon the retina.

Factors which influence the pattern of impulses at the cortex may be of three general types: (1) the size, number, and separation of retinal elements; (2) the frequency of impulses produced by each element; and (3) possible interaction between elements. These factors can hardly be discussed separately.

The impulse frequency is presumed to be a function of several physiological factors, including the intensity of the exciting stimulus, the state of fatigue or recovery of the particular neural unit involved, and possibly the suddenness with which the stimulus impinges upon the receptor. In general, interrupted light assures more opportunity for all elements to be non-refractory when the stimuli are applied, and this effect is, in turn, related to the natural periodicity of neural recovery.

Furthermore, as has been suggested by Marshall and Talbot (11), physiological nystagmus may operate to assure availability of the visual pathway at any given moment in a fatiguable system. These authors postulate a dynamic interrelationship between the steepness of a moving gradient of intensity impressed on a receptor and the width of that receptor. It is usually found that the steepest slope of the intensity distribution is the same width as the receptor element (11). This means that as the moving image shifts from receptor to receptor, the suddenness, as well as the amount, of change in intensity operates to enhance the frequency of neural impulses.

When we interrupt the light at frequent intervals we further increase the suddenness of photic change. The movement of the distribution of intensities across the receptors is no longer continuous. It is spatially discontinuous, since the image of an edge or border appears first at one position and then at a different position on the retina, while the receptors in between are not stimulated at all. The change is no longer one from very little intensity to much intensity, but from no intensity to total intensity.

An examination of Fig. 5 may suggest a further peaking mechanism which

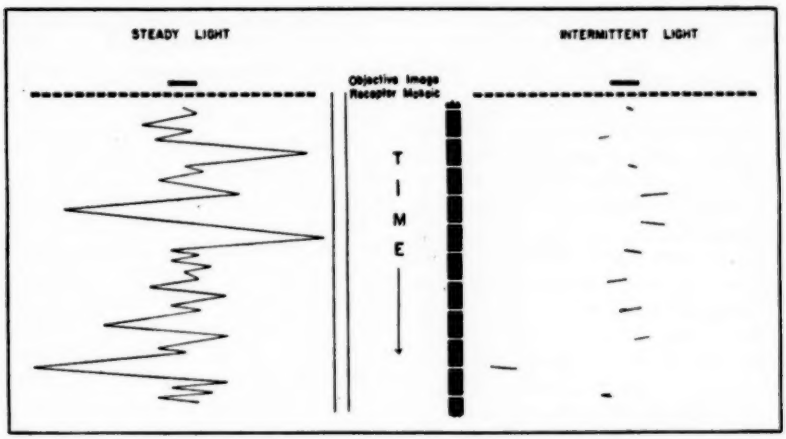

FIG. 5. A diagrammatic representation of the effect of interrupting the light on the stimulation of retinal receptors. The situation with steady light is schematically illustrated on the left; on the right the situation with a light-time fraction of 0.10 is shown. The eye-movements are the same in both cases.

might take place with interrupted light. Although the relative frequency with which the moving image impinges on each retinal element is the same for both interrupted and steady light, the result in terms of impulse frequency is likely to be different. For with steady light, those cones farthest from the objective illuminated point will be illuminated least frequently. For this reason, the neural units farthest from the center of the intensity distribution will be non-refractory for a large portion of the time, whereas the cones at the center of the intensity distribution will be stimulated most often, because of the statistical nature of the distribution of eye-movements; hence they will be refractory during a greater portion of the time. This is the situation as it exists with steady light.

When the light is intermittent, the relative frequency with which the image impinges upon the cones at the center of the intensity distribution is the same as in the case of the steady light. But because of the relatively long dark intervals between stimulations the 'central' receptors of the intensity distribution will more often be non-refractory. The net result of this will be to increase relatively the number of impulses from the neural units in the center of the fluctuating distribution of intensities, without increasing proportionately the number of impulses from the more distant neural units. Therefore, the overall 'averaged' distribution will show more peaking at the center of the intensity distribution when the light is interrupted than when it is steady. Similar effects probably occur at higher levels of the visual projection system, and the total result is probably modified and augmented by temporal summation.

At the same time that the relative frequency of neural impulses from the central receptors is increased, the total number of impulses delivered may be decreased, or may remain the same, because of a decrease in opportunity for spatial summation. Since the chance for adjacent receptors to be stimulated simultaneously or in immediate succession is decreased, there should be a resulting decrease in spatial summation, and with it, a decrease in apparent brightness. This would explain why resolvability could be increased without an increase in brightness. And since these two opposing factors would seldom

be expected to balance exactly, it would explain why the resolution contours are not straight lines but exhibit maxima and minima.

Whether the mechanism suggested above actually operates has yet to be determined experimentally. But it does seem reasonable to believe that any explanation of the results obtained in this study must be stated in terms of some peaking mechanism operating between the retina and the cortex. Any theory of visual acuity which postulates a one-to-one relation between photochemical changes at the retina and visual response does not seem adequate to account for the obtained data.

Summary

In order to test the hypothesis that visual acuity is a form of brightness discrimination, a situation was sought in which visual acuity and brightness discrimination might be shown to have a different functional relation to the same independent variable. By using interrupted, rather than steady light, and measuring the intensity required for resolution as a function of the light-time fraction in the interruption cycle, such a situation was found. The resolution contours obtained in this manner lay significantly below the brightness contour, as expressed by the Talbot-Plateau Law. Extensive results were obtained from three *O*s, under several different conditions of image size, retinal location, flash frequency, and wave-length composition.

It was therefore concluded that visual acuity cannot be a form of brightness discrimination, although the two processes may be related. Specifically, this means that the quantitative generalizations applicable in one case need not necessarily be applicable in the other. A mechanism has been suggested to account for the obtained data. Peaking of the image may occur because the retinal elements in the center of the fluctuating distribution of intensities are non-refractory for a relatively greater portion of the time when the light is interrupted than when it is steady, resulting in an increase in the relative porportion of neural impulses coming from this central portion. The data cannot be accounted for at a strictly retinal level; any explanation must take account of neural processes central to the retina.

(Manuscript received August 16, 1948)

References

1. Bartley, S. H. Subjective brightness in relation to flash rate and the light-dark ratio. *J. exp. Psychol.*, 1942, **31**, 380–92.
2. Bartley, S. H. *Vision.* New York: Van Nostrand, 1941.
3. Crozier, W. J., & Holway, A. H. Theory and measurement of visual mechanisms: I. A visual discriminometer. II. Threshold stimulus intensity and retinal position. *J. gen. Physiol.*, 1939, **22**, 341–64.
4. Crozier, W. J., & Wolf, E. Theory and measurement of visual mechanisms: IV. Critical intensities for flicker, monocular and binocular. *J. gen. Physiol.*, 1941, **24**, 625–33.
5. Crozier, W. J., & Wolf, E. Theory and measurement of visual mechanisms: XI. On flicker with subdivided fields. *J. gen. Physiol.*, 1944, **27**, 401–32.
6. Hecht, S. A quantitative basis for the relation between visual acuity and illumination. *Proc. nat. Acad. Sci., Wash.*, 1927, **13**, 569–74.
7. Hecht, S. The nature of the photoreceptor process. In: Murchison (Ed.), *Handbook of gen. exp. Psychol.*, Worcester: Clark Univ. Press, 1934. Pp. 704–828.
8. Hecht, S., & Mintz, E. The visibility of single lines at various illuminations and the retinal basis of visual resolution. *J. gen. Physiol.*, 1939, **22**, 593–612.
9. Hecht, S., & Wald, G. The visual acuity and intensity discrimination of *Drosophila*. *J. gen. Physiol.*, 1934, **17**, 517–47.

10. HECHT, S., & WOLF, E. The visual acuity of the honey bee. *J. gen. Physiol.*, 1929, **7**, 727–60.

11. MARSHALL, W. H., & TALBOT, S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In: KLÜVER (Ed.), *Biol. Symp.*, 1942, **7**, 117–64.

12. PLATEAU, M. Sur un principe de la photometrie. *Bull. Acad. Roy. Soc., Brussels*, 1835, **2**, 52–60. Also: note in **3**, 89.

13. SHLAER, S. The relation between visual acuity and illumination. *J. gen. Physiol.*, 1937, **21**, 165–88.

14. TALBOT, W. H. F. Experiments on light. *Phil. Mag.*, 1834, **5**, 321–34.

RESPONSE LATENCY AS A FUNCTION OF THE AMOUNT OF REINFORCEMENT[1]

BY DAVID ZEAMAN[2]

Columbia University

Introduction

The evidence necessary to support the framework of a quantitative, reinforcement theory of behavior includes that provided by the functional relationships established between measures of response strength and the various aspects of the stimulating surround that act as parameters of the basic processes of acquisition and extinction.

Of considerable systematic importance is the parameter of amount, or quantity, of positive reinforcement, a variable which is capable of generating families of learning and extinction curves. Before this parameter can be made the independent variable in a simple learning experiment, it is necessary to specify the measure of response strength chosen, since there is some indication that there are as many equations for acquisition and extinction as there are combinations of different measures of strength and different estimates of central tendency of these measures (3, 12, 13).

Restricting consideration to the operant, striated muscular response, the specifications of strength chosen by previous investigators fall roughly into four categories determined by the dimension or unit of strength employed. Each of these has been used in experiments involving amount of positive reinforcement as one of the independent variables. These are: latency of response in units of time (15, 16); number or frequency[3] of response (22), which includes resistance to extinction (5); response force or work in physical units (6); and speed of response in units of time-distance ratio (1, 11).

To meet adequately the demands of a quantitative theory, the results of the above investigations would have been reported in the form of functional relationships between amount of reinforcement and the value of one or more constants in the acquisition and extinction equations. Preferably, the data for such equations should be obtained by use of one of the more conventional experimental techniques such as those employing the lever-pressing (20), or runway situation (9). Although the actual results presented do not take this form, some reasonable inferences can usually be drawn as to the nature of the effect of amount of reward on the equations of the basic processes.

Hull, in his *Principles of behavior*, (13), reviewed the admittedly inadequate evidence existing at that time and put forth the tentative hypothesis that amount of reinforcement affected the learning curve asymptote, or equation constant representing the final limit, rather than the constant controlling the rate-of-approach to that asymptote. This hypothesis received some degree of confirmation through the experimentation of Crespi (1),[4] who used speed of

[1] A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University. The research was carried out under the sponsorship of Professors F. S. Keller, W. N. Schoenfeld, and C. J. Warden.

[2] Now at Brown University.

[3] There is no evidence reported on the effect of amount of reinforcement on a time-rate measure of response strength such as that employed by Skinner (20).

[4] An extensive and detailed survey of background material pertinent to the general problem has been given by Crespi (1). The present experiment was performed without knowledge of this study by Crespi.

response as an index of strength on a 20-ft. runway. Crespi found that larger amounts of food reinforcement increased the maximum running speeds built up in rats by long, well-spaced practice. Empirical learning curves showed the rate-of-approach to these maxima for different conditions of amount to remain approximately constant, as well as could be determined by an inspection of the number of trials taken for each curve to reach its asymptote. Crespi considered his results to be in general agreement with Hull's tentative hypothesis despite the fact that the hypothesis concerned, not empirical learning curves, but the relationship between the number of reinforcements and the intervening variable, habit strength ($_sH_R$). Only in the event that habit strength was linearly related to the dependent variable of Crespi's empirical learning curves—running speed—could the hypothesis be said to predict logically the experimental results. Actually, in Hull's system, under conditions of constant drive, habit strength is proportional to a second construct, reaction evocation potential ($_sE_R$), which is anchored on the level of observation to four measures of response strength, none of which includes running speed.[5] If one of these measures, amplitude or magnitude of reaction (A), postulated to refer only to autonomically mediated responses, were extended to include striate muscular reactions as well, and if finally there were postulated a linear relation between reaction potential and magnitude of reaction, then Hull's hypothesis would predict, and in turn be strengthened by, Crespi's results. Application of the habit-strength quantificational techniques recently proposed by Hull, Felsinger, Gladstone, and Yamaguchi (12) might obviate the necessity of the last special assumption.

Crespi (1) also reported on the effect of variation of amount of reinforcement on already established learning curve asymptotes. Increments in amount of reinforcement were found to shift upward the level of performance, whereas decrements in amount led to extinction effects, a lowering of the level of performance. Changes in amount, however, were accompanied by changes in level of performance greater than would be predicted from a knowledge of the asymptotes of animals trained on single amounts. There were, in short, contrast effects. Since the limits of response strength were found to be related not only to present amount of reinforcement, but also to the past history of amount, Crespi constructed cognitive equivalents, in the rat, of that past history—an 'expectancy' or 'eagerness' for a given amount; and in order to account for the positive and negative contrast effects, there were constructed emotional drive states of 'elation' and 'depression' produced by non-fulfillment of expectancies. The status of these constructs will be examined in greater detail in a later section of this paper. It may be stated here, in summary, that Crespi's results clearly established amount of reinforcement as an effective parameter of the process of acquisition, in control specifically of the asymptote of the empirical learning function when running speed was the measure of response strength.

Part of the evidence that led to Hull's amount-of-reinforcement hypothesis came from researches in which various response strength measures were used, but until more recently a very important measure, latency, had not been subjected to investigation with amount or reinforcement as a variable.

Lawrence and Miller (16), using a runway situation, found that groups of rats receiving .16 gm. were superior in their latency scores to those receiving .04 gm. during acquisition and extinction, although the differences between the two groups were not statistically significant. Kimble (15), in an unpublished study, found that a change from .3 to .04 gm. of food reward was sufficient to alter the level of performance as measured by latency scores on a runway. A composite variable of amount of reinforcement and temporal spacing of trials was re-

[5] Hull's own use of running speed as a measure of $_sE_R$ is also not consistent with his postulates (13, p. 344).

ported by Gagné (7) to relate systematically to runway latencies, but no conclusion as to the effect of amount alone could be drawn. This assembly of experimental data does not provide a well-defined picture of the role of amount of reinforcement in determining the course of the behavioral processes when latency is the index of performance.

If latencies were a characteristically poor measure, this deficiency would not be of great consequence. But latency, on the contrary, has proven to be a fruitful measure, for a number of reasons which may profitably be reviewed here. Latency has been used in conventional types of experimental situations such as the lever-pressing, jumping-stand, and runway (3, 4, 9, 7, 8, 17, 18, 19, 21), to demonstrate successfully a wide range of behavioral phenomena. As has been pointed out by Graham and Gagné (9), latency quantifies in units of a known physical dimension, and thus facilitates comparison with the data of related disciplines. This measure, when expressed in logarithmic form, tends to give a normal distribution of scores. With means and medians roughly coincident, the number of quantitative laws necessary to describe the same process is thus reduced. For these reasons, in addition to a practical consideration of ease of recording, latency ranks high in importance as a quantitative index of response strength, and the sensitivity of this measure to such a major stimulus variable as amount of reinforcement takes on significance.

The purpose of the present investigation lies in the determination of the functional relationship between the amount of positive reinforcement and one or more of the constants of the empirical acquisition and extinction equations containing latency of response as dependent variable. An auxiliary purpose seeks the answer to the question, "What is the effect of a change in amount of reinforcement on latencies late in the stages of acquisition?"

Apparatus

The apparatus employed in this experiment was a modified form of that devised by Graham and Gagné (9).[6] It consisted of a simple, straight runway 1¼ in. wide and 3 ft. long, elevated 3 ft., with a small platform (7 × 10 in.) at each end. A starting box with inside dimensions (7¾ × 6½ × 6 in.), open in front and covered on top with wire screen, was placed on one platform; a food box of the same dimensions was placed on the other. Small doors of the guillotine type separated the platforms from the runway. These could be quietly raised and lowered by the experimenter, thus controlling the animal's access to the runway from either box. The back of the starting box was hinged to permit entrance of the rat prior to each trial. The back of the food box was provided with a handle by means of which the experimenter could take the box from its platform to remove the rat at the termination of each trial.

With the exception of the platforms and boxes, the whole apparatus was contained within a wooden frame covered with cheesecloth. The runway was illuminated from within the framework (illuminometer reading of 1.38 ml. from the runway). Since the experimental room was darkened, the cheesecloth served as a relatively effective one-way screen.

The exposed surfaces of the apparatus, including the starting box, were painted a flat white with the exception of the food box, which was painted a flat black to clearly differentiate it from the starting box.

Measurements of latency and of running time were taken by means of two electric .01-sec. timers controlled by manual switches located on the frame of the apparatus.

In another part of the experimental room was an elevated practice runway of the same dimensions as the experimental runway, but not enclosed within a framework and not provided with platforms or boxes at either end.

Animals and Procedure

Animals.—101 male white rats of Wistar strain were used in this experiment. Of these, 96 contributed to the data presented, 3 having been excluded for failure to leave the starting box on the first trial within 5 min., and 2 having been discarded after accidental injury.

Table I describes the basis for distribution of the animals into three groups. The ages and

[6] The apparatus was constructed by Margaret Raben for use in an investigation of stimulus generalization.

TABLE I

CHARACTERISTICS OF THREE EXPERIMENTAL GROUPS

	N	Age (days)	Average Weight (ounces)	Previous Training	Experimental Treatment
Group I	40	80 to 120	8.0	None	Acquisition with varying amounts of reinforcement; and extinction
Group II	40	90 to 110	8.6	None	Acq; further acq. with change in amount of reinforcement; and extinction
Group III	16	Approx. 100	8.0	Lever pressing, acq. and ext.	Same as Group II

weights tabulated are those at the beginning of the experiment. An average weight increase of 0.4 oz. was recorded at the end of approximately one month's experimentation. Litter relationships were not known, and were assumed to be randomly distributed among all groupings. Groups II and III were run in the same daily sessions during the month of August; those in Group I during July.

Preliminary training.—All animals were given 10 daily sessions of preliminary training on the practice runway, during which time they were maintained on a schedule of one daily feeding of 2.4 gm. of 'process' cheese and 4.6 gm. of ground Purina dog checkers given after the daily session. Each rat was allowed to run without reinforcement on the practice runway for a minimum of three min., extended until at least two complete traversals had been made. The number of traversals of each rat during the three min. was recorded in order to serve as a rough index of unconditioned activity level and subsequently as a basis for equating groups of animals prior to the experiment proper. Following this practice run by some variable time period within an hour, the animal was habituated for several minutes to the starting and goal boxes which had been removed from the experimental runway. On the last three days of preliminary training, a few small pieces of cheese were placed within the small dish inside the goal box. While this procedure would tend to cause the goal box to become a secondary reinforcing agent present at the end of the first acquisition trial—a superfluous condition for the processes to be studied—, it had the advantage of also ensuring immediate primary (food) reinforcement for those animals who were to receive very small pieces of cheese. These animals might otherwise encounter their reinforcement in the goal box only after considerable delay. At the termination of the preliminary training the animals of Group I were divided into subgroups of 10, each being equated for weight and activity level on the practice runway, and each assigned a different amount of reinforcement, .2, .4, .8, or 1.6 gm. to start acquisition. Group III was divided into similarly equated halves, one to receive .05 gm., the other 2.4 gm. on each trial of acquisition. Group II was given identical preliminary training, but was not subdivided into smaller groups before starting acquisition on .60 gm. of reinforcement.

The average deviation of all cheese weight measurements varied from 5 percent for the heaviest pieces to 10 percent for the lightest.

Experimental procedure.—Any experiment designed to isolate the effects of different quantities of reinforcement must separate and control changes of drive state and spacing of trials, since these factors are known to be relevant, and tend to vary concurrently with changes in the amount of reinforcement. Larger amounts of food take longer to eat and decrease the drive state to a greater extent than do smaller amounts. Since it was desired to make a trial-by-trial comparison between groups rewarded with different weights of cheese, it was necessary that rats of all groups enter each trial with the same drive and at a constant interval after the preceding trial. The only solution lay in the radical spacing of trials. One trial a day brings each rat into the experimental situation with approximately 23 hours' food drive and at an inter-trial interval of 24 hours. While eating time did not affect temporal spacing of trials, it was significant in setting a practical upper limit on the amount of reinforcement. Doubling the largest amount of 2.4 gm. would have doubled the length of the daily experimental session, making it difficult to deal with the same numbers of animals.

Except for the modifications required by the independent variable and long inter-trial periods, the specific details of procedure were nearly the same as those usually employed with this runway. The animal was carried manually from the living cage to the starting box where it remained for a minimum of 15 sec. When the animal faced the door leading to the runway, the door was opened, and the latency of the response of leaving the box (disregarding the rat's tail) was measured. The difference in time between the exit from the starting box and the entrance into the goal box, the running time, was also measured.[7] The door to the goal box was closed gently behind the rat, and he was allowed just sufficient time to eat the quantity of cheese allotted him before his removal to the living cage. At a variable time between 40 and 60 min. later, the animal received 4.6 gm. of ground dog chow, and an amount of cheese equal to the difference between the particular amount received as reinforcement and 2.4 gm. In this manner the diet for all animals was kept constant.

Acquisition with varying amounts of reinforcement.—The four groups receiving .2, .4, .8, and 1.6 gm. were given 18 daily acquisition trials, at the end of which the average log latencies for all groups had not differed significantly for at least the last 6 trials. The .05- and 2.4-gm. groups also achieved asymptotic latencies within 18 trials, but were given one additional day's training because it was not expedient on that day to change them to a new schedule. The 40 animals comprising the .60-gm. group were given just 18 learning trials.

Extinction.—For 10 trials on the day following acquisition, the animals in the .2-, .4-, .8-, and 1.6-gm. groups were changed to zero amount of reinforcement, and thus underwent extinction. The need for one-trial-a-day spacing no longer being present, the inter-trial interval in extinction was made as low as possible to hasten the process. A massing interval of 20 sec. was employed. It had been intended to use as a criterion of complete extinction the conventional 3 or 5 min. of lack of response, but so many animals failed to meet either of these criteria within 50 trials that the extinction period was finally set at 10 trials. Further decrement in response magnitude beyond the first 10 extinction trials was generally not great in those cases in which observations were made. Furthermore, the fact that all animals ran at least 10 extinction trials, but not all ran 11 trials, makes the choice of the number 10 somewhat less arbitrary.

Further acquisition with a change in amount of reinforcement; and extinction.—On the day following the last acquisition trial, the two groups of eight animals receiving .05 and 2.4 gm. had their respective amounts of reinforcement reversed, and were given eight further daily trials with the opposite amounts.

Also on the day following the last acquisition trial the 40 rats in the .60 group were subdivided into five groups of eight animals, equated for weight and for the average log latency on the last six days of acquisition. Each group was then assigned one of the following amounts, .05, .20, .60, 1.20, or 2.40 gm., and was run for eight further daily trials under the new conditions of reinforcement.

An extinction procedure identical to that described above was carried out on all animals in these change-in-amount groups.

Results and Discussion

Acquisition.—Fig. 1 shows the family of acquisition curves generated by the parameter of amount of reinforcement. The points through which the curves have been drawn represent the average log latent period on each trial for seven groups of rats running for .05, .20, .40, .60, .80, 1.60, and 2.40 gm. of reinforcement.[8] The upper set of curves in Fig. 1 permits an inspection of the goodness of fit of the empirical points to the curves. The curves have been fitted to the data by means of a slightly modified form of the equation found by Graham and Gagné (9) to describe the acquisition process The

[7] The running-time data will not be presented. The running-times proved to be less sensitive to variation of the parameter of amount of reinforcement than the corresponding latency measures. The direction of change was always the same as that of the latency measure, but the magnitude was not as significant.

[8] For the numerical values of these average log latent periods, the standard errors of the log values at each trial, and similar tables for all graphical material presented in this paper, order Document 2650 from American Documentation Institute, 1719 N Street, N.W., Washington 6, D.C., remitting $0.50 for microfilm (images one in. high on standard 35 mm. motion picture film) or $0.50 for photocopies (6 × 8 in.) readable without optical aid.

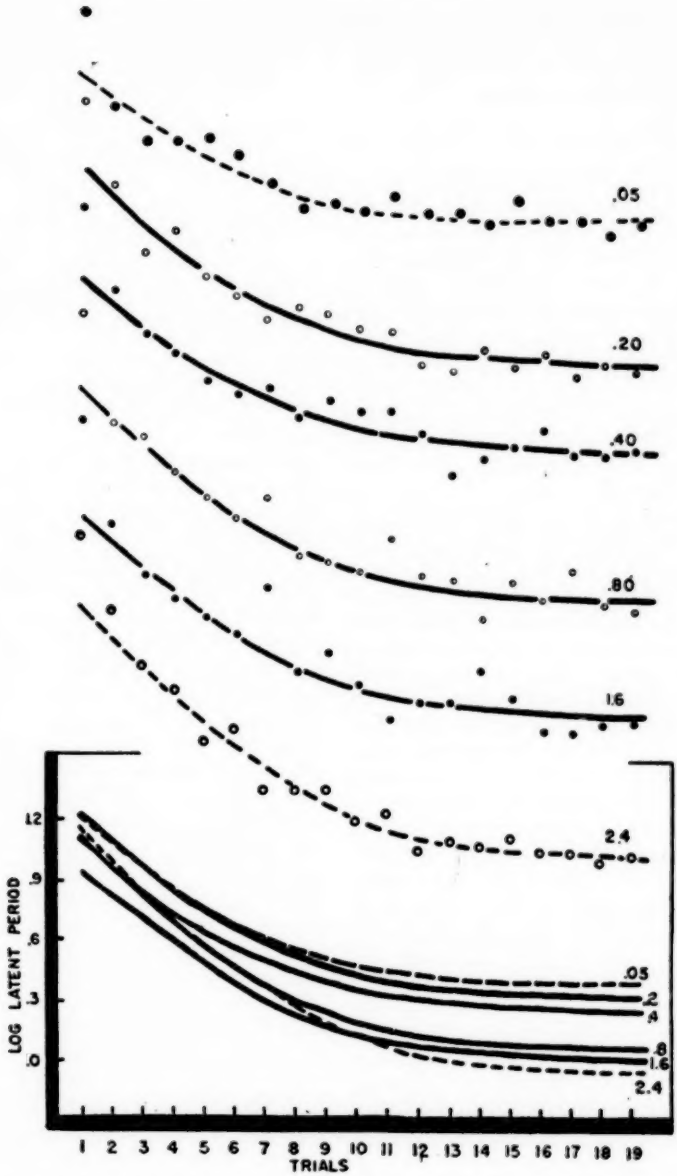

FIG. 1. The lower set of curves describes the course of acquisition for each of six different amounts of reinforcement. The upper set of curves is the same set separated on the ordinate to facilitate inspection of the experimental points of each amount-group.

modification[9] consists of substituting the expression $(n - 1)$, the number of reinforced trials, in the original equation which then becomes

$$P = (P_m - P_\infty)C^{-n+1} + P_\infty \quad (1)$$

where P is the latent period at any trial; P_m, the 'true' first trial latency; P_∞, the latency after many trials, i.e., the asymptote; and C, a constant presumed to have spacing of trials as one of its determinants (**7, 9**). By 'true' first trial latency is meant that latency which the animals would assume if they were allowed to run for a number of trials with no reinforcement. This equation has the advantage of placing P_m at trial 1 (first reinforcement), where it may be compared with the empirical latency value on that trial. Mote and Finger (**17**) have demonstrated that the best estimate of unconditioned operant level was in the range of .44 log latent period units below that observed on the first trial. Fig. 1 reveals that for every group the calculated P_m fell below the corresponding empirical points, the average decrement being .37. P_m was calculated, together with C, by transposing equation (1) to linear form, a plot of log $(P - P_\infty)$ against $(n + 1)$ conforming approximately to a straight line with a negative slope equal to the antilog of C and a y-intercept equal to the antilog of $(P_m - P_\infty)$.[10] The value of the asymptote, P_∞, in those calculations was obtained for each amount-group by averaging the mean log latencies of the last six trials of acquisition (trials 14–19). Table II presents constants of the acquisition equation for each of the quantities of reinforcement except those in Group II, the .60-gm. amount-group, which proved to be not comparable. It had been intended that the 40 animals in the .60 group would be comparable to the previous experimental animals in all relevant respects. To this end, a group of animals were obtained, equated to the previous animals for the center of the age range. This criterion did not prove sufficient to provide the near identity of the two groups for weight and activity level on the practice runway. Of these two variables, weight is undoubtedly not significant, since it did not correlate with the asymptote of the individual animals within the .60 group. The difference in activity level can possibly be ascribed to litter differences which were not known. This difficulty was not encountered with the animals in the other amount groups who were matched for activity level prior to experimentation.

TABLE II

VALUES OF CONSTANTS * IN ACQUISITION AND EXTINCTION EQUATIONS UNDER VARYING CONDITIONS OF AMOUNT OF REINFORCEMENT

Amount in grams	Acquisition			Extinction		
	P_∞	C	P_m	P_f	c	P_1
.05	2.30	1.45	16.2			
.20	1.90	1.45	15.7	.20	1.50	1.90
.40	1.67	1.40	10.7	.20	1.38	1.67
.80	1.07	1.45	11.8	.20	1.20	1.07
1.60	.96	1.45	7.9	.20	1.10	.96
2.40	.80	1.45	14.6			

* These constants are expressed in units of seconds.

Fig. 2 depicts the course of acquisition for the 40 animals in the .60-gm. group. Two equation forms are fitted to the data, an exponential (equation 1) and a power function of the type used by Felsinger, Gladstone, Yamaguchi, and Hull (**3**) to describe the

[9] This modification does not greatly change the shape or position of the curve except for the very early trials in acquisition, the least reliable portion of the curve.

[10] As determined by the method of averages.

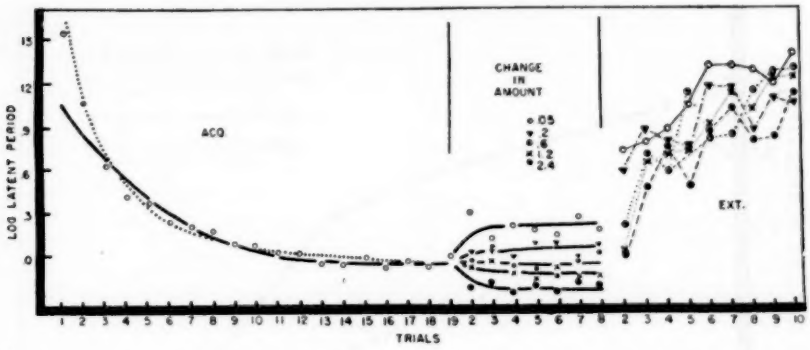

FIG. 2. Acquisition data of the 40 animals in the .60-gm. group, fitted by two types of equation forms, an exponential and a power function. The exponential equation (solid line) is:

$$\text{Latency} = 15.2\ (1.54)^{-n+1} + .91,$$

and the power function (dotted line) is:

$$\text{Latency} = 50\ (n)^{-2.3} + .91,$$

where n is the number of reinforcements. Eight further acquisition trials under changed amounts of reinforcement are represented by curves drawn by inspection through the experimental points. The extinction process is represented by the experimental points connected by straight lines.

empirical latency learning curve obtained from the rat in a bar-press situation. The power function fits the data for the early trials better than does the exponential equation, and hence must be regarded as the better description of the empirical latency learning curve. On the other hand, the deviation of the exponential curve from the first point cannot be evaluated as simple error, since the deviation has testable consequences involving the independent measurement of the unconditioned operant level. The choice between the two forms of equations can be left open to theoretical considerations.

The effect of amount of reinforcement on the process of acquisition is seen most readily in Table II, which lists the constants of the equations for comparable amount groups, and in Fig. 3, which graphically relates the independent and one of the dependent variables of this investigation. Of these three constants, only P_∞, the asymptotic latency, undergoes systematic variation with the amount of reinforcement. Consequently, this is plotted in Fig. 3 as a function of the log gram weight of reinforcement. The resultant curve is negative in slope and tends to be slightly positively accelerated over the weight range tested.

The constant C is relatively invariant for all conditions of amount; and while P_∞—the predicted starting point of conditioning—does vary considerably, its variation is not systematic.[11]

[11] Many replications of the present experiment would provide groups of animals more nearly equated for the 'true' first trial latency (P_∞). The reduction in variation of P_∞ either experimentally through replication or by means of some suitable statistical manipulation of the entire functions as has been suggested by Professor I. Lorge, might alter the relationships found to hold between amount of reinforcement and the other parameters, C and P_∞. The seven first-trial, experimentally observed, latencies contained no significant differences.

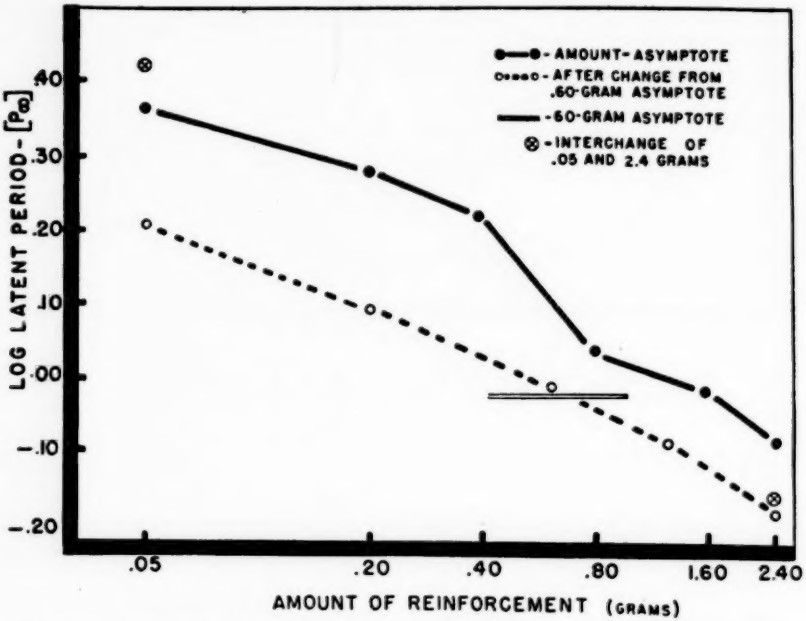

FIG. 3. The relation between the quantity of food reinforcement and the final latency asymptotes established by training with constant amounts (solid lines), by changing the amounts from an intermediate, .60 gm. level (broken lines), and by interchanging the smallest and largest amounts of reinforcement (crossed circles).

Whether P_∞, the mean of trials 14–19, is justifiably looked upon as an asymptote has been statistically tested.[12] A null hypothesis was set up that the criterion variable P_∞ was unaffected by further increases in the trial variable; or, in other terms, that the series of latency means from trial 14 through trial 19 was a horizontal straight line. *F*-tests were performed with a 'trials × animals interaction' variance used as the measure of experimental error. None of the probability values of the *F*'s obtained approached the .05 level of significance, thus permitting the null hypothesis to be retained. In reference to latencies, the term 'asymptote' throughout this paper is statistically defined as meeting this *F*-test for horizontal linearity. It does not mean a value obtained by extrapolation to the point of infinite trials.

The reliabilities of the asymptote differences produced by the variation of amount of reinforcement were determined by *t*-tests. These proved to be very insensitive tests because of the large variance within groups. A generalization drawn from an examination of individual *t*-tests sets the magnitude of increase of amount of reinforcement necessary to produce a significant latency decrease, at a minimum of a fourfold change. This

[12] The choice of trial 14 as the point at which asymptotes begin was determined by an inspection of the present curves and other similar runway latency curves in the literature. These all indicate that latency changes do tend to occur before but not after the 14th trial.

corroborates a similar finding by Crespi (1) for a different measure of response strength.

Another statistical test of the significance of the differences between amount-asymptotes proved to be more sensitive. It was noted that the *t*-test frequently showed two sets of asymptotic latencies for two amounts to be not significantly different, when the individual points of one set always fell above the points of the other for the six asymptotic trials. If the assumption is made that at each trial there are two points randomly drawn from the same population, the probability of the point from the one amount-group falling above that of the other is equal to $\frac{1}{2}$. The usual method of combining probabilities based on the binomial expansion will give the chance frequency of any combination of 'aboves' and 'belows,' with no assumption made as to the distribution characteristics of the point population sampled at each trial.[13]

TABLE III

PROBABILITIES OF THE CHANCE OCCURRENCE OF THE OBSERVED COMBINATIONS

Amount (grams)	2.40	1.60	.80	.40	.20
.05	.02	.02	.02	.02	.02
.20	.02	.02	.02	.11	
.40	.02	.11	.02		
.80	.02	.11			
1.60	.34				

The chance frequency of obtaining six consecutive 'aboves' for a given amount-group is 1/64, or two percent. The smaller amount-groups' mean latency always tended to be above that of the larger. Table III contains the probabilities that the observed frequencies of being above were due to chance.

Eleven out of the 15 possible asymptotic differences are seen to be significant at the 2 percent level of confidence. Application of the same test to the entire learning curves, exclusive of the first trial, increases the number of significant differences to 13 out of 15. The failure reported by Lawrence and Miller (16) to get significant differences in latencies between their .04- and .16-gm. groups is paralleled in the present experiment when the *t*-test is applied to the difference between the asymptotes of the .05 and .20 groups. The simple combined probabilities test, however, reveals that the difference between the .05- and .20-gm. groups is significant at the two percent level.

Extinction.—The systematic weakening of the strength of response during the process of extinction is shown in Fig. 4. A sharp increase in the latencies of all groups can be observed during the first few trials of extinction, followed by a gradual sloping off of the rate of rise, until at trial 10 the curves appear to be approaching a common level. Approximately half of the animals in each group were carried beyond the tenth trial in an unsuccessful attempt to find the number of trials necessary to reach a criterion of three minutes of no response, and for all these animals the latency variation after trial 10 centered about a common geometric mean of 20 sec. This value of 20 sec. was set equal to P_f, a constant representing the final level of response strength in the equation

$$P = (P_1 - P_f)c^{-p+1} + P_f, \qquad (2)$$

where P is the latency at any trial; P_1, the first extinction trial latency (and last acquisition trial); c, an empirical constant; and p, the ordinal

[13] Grant (10) presents a convenient formula for this type of test.

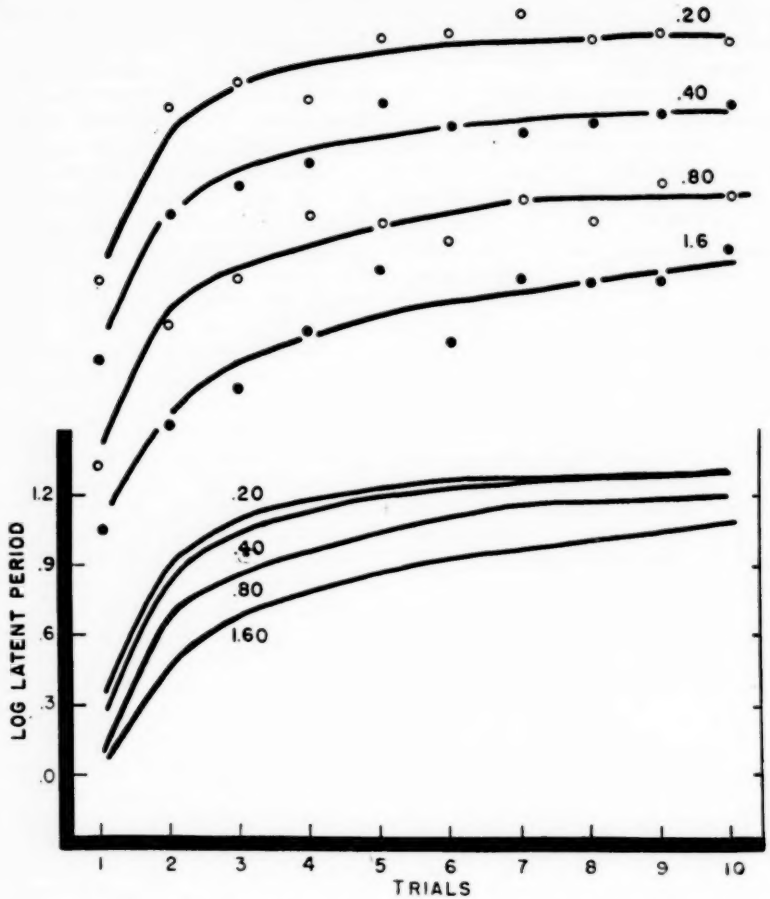

FIG. 4. Extinction of each of four amount-groups. The lower set of curves has been plotted from values calculated from equation (2) and the constants of Table III. The upper set of curves is the same set elevated on the ordinate to prevent overlapping of the experimental points which have been included.

number of the extinction trial. This equation has previously been found to be descriptive of the extinction process as measured by runway latency (9). A curve-fitting procedure was employed identical to that used for the acquisition equations. Fig. 4 shows the goodness-of-fit of the computed curves to the experimental points. The values of the constants contained in the equations of the different amount groups are listed in Table II.

To be contrasted with the fact that amount of reinforcement caused systematic variation of the acquisition equation constant, P_∞—the asymptote or final level—is the observation

that the final level in extinction remains invariant while c, the constant controlling rate-of-approach to the final level, does functionally relate to the amount parameter. Since resistance to extinction may be indicated by either slower rate of approach to a final limit, or by a lower final limit, or both, the variation in c may be interpreted as meaning that greater resistance to extinction is associated with greater amounts of reinforcement.

The significances of the differences between the four sets of extinction points have been tested by extending the probability technique described above to apply to the nine extinction trials. The differences between all amount groups except between the .20- and .40-gm. groups are significant on at least the two percent level.

Change in amount.—The effect of changes of reinforcement on the asymptotic latencies reached by animals trained on an intermediate amount of reward of .60 gm. is represented graphically in the section of Fig. 2 designated 'change in amount.' The course of change is shown for the five equated groups during the eight additional trials under the amounts specified.

The one group maintained on .60 gm. as a control group deviated very little from the .60 asymptote latency of −.02 log units. The mean latency of the last six of the eight additional trials for this group was −.01 log unit. The remaining four groups underwent changes in latency directly related to the change in amount of reinforcement. The last six trials for these groups met the F-test for horizontal linearity with mean latency asymptotes for the .05-, .20-, 1.2- and 2.4-gm. groups equal, respectively, to .21, .05, −.09, and −.19 log units. Here, as with animals trained with no change in amount, the t-tests of the significance of the asymptote differences show that a minimum of a fourfold change in amount of reinforcement is required to produce a significant asymptote difference (five percent level). At the lower amount levels, as great as a twelvefold change is necessary for the effect to take place.

Again, the simple combined probabilities test of significance proved the more sensitive, and brought out significant differences on the one percent level for 8 of the 10 possible comparisons. The remaining two comparisons bordered on significance at the six percent level.

A comparison of the after-change asymptotes for each amount with those obtained by the no-change procedure is made most clearly in Fig. 3. The shapes of the two functions are similar, both being negative in slope and positively accelerated. A marked difference is present, however, in the position of the two curves with respect to the latency ordinate. The difference is consistent with the measurement taken during the preliminary training which showed the 40 animals originally in the .60-gm. group to be faster-moving animals. If this, presumably native, speed or activity difference is cancelled graphically by moving the change in amount function upward, by approximately .13 log latency units, there appear no 'contrast' effects, i.e., the largest changes in amount do not produce changes in latency asymptotes greater than that predicted by the simple learning curve asymptotes under different amounts. This negative evidence does not weigh heavily, inasmuch as the maximum changes in amount (12-fold decrement and a 4-fold increment) might not have been great enough to show the effect.

The interchange of amount of reward of the .05- and 2.4-gm. groups provides a fiftyfold change of amount, and hence a more critical test for the presence of contrast effects. The results of this test are presented in Fig. 5. Following the establishment of the horizontal performance levels at .36 log units for the .05-gm. group and at −.08 log units for the 2.4-gm. group, the reversal of amounts for these two groups on the 20th day's trial caused an immediate, significant reversal of the latency magnitudes, measurable on the next day's trial and persistent throughout the remaining six days of changed amount. Although every trial mean difference between the .05- and 2.4-groups is significant at the one percent level, the comparison which must be made to test for the contrast effect is between the asymptote before and after change for each of the two amounts. An extrapolation of the pre-change asymptote into the post-change trials has been made in Fig. 5, which reduces the contrast effect to an area between the pre- and post-change asymptotes. The significance of this effect can be tested by setting up the assumption that there is no difference between the populations of pre- and post-change asymptote latencies, from which it follows that the chances of a post-change latency value falling above the extrapolated pre-change mean are equal to the chances of it falling below. An examination of Fig. 5 reveals that four of the six trial points for the .05 group fell above the predicted mean, and that the entire six points of the 2.4 group fell below the expected mean. The probabilities of these frequencies, given the assumptions of equiprobability, are 7/64

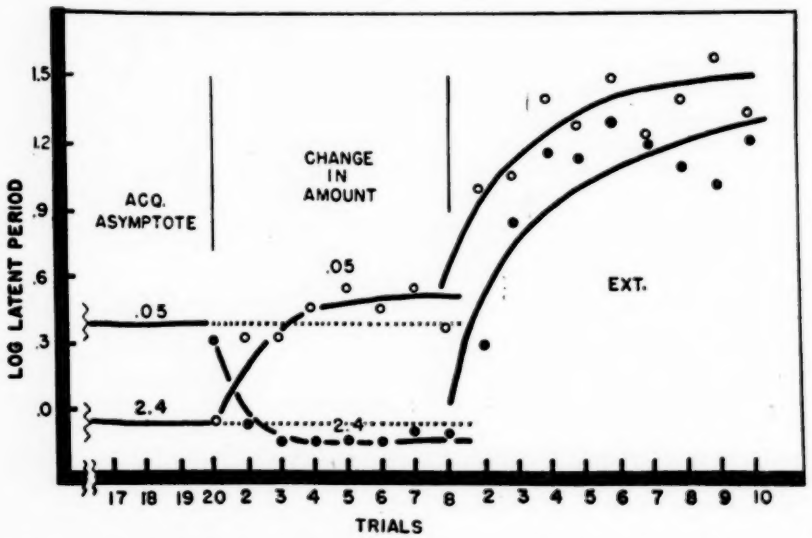

FIG. 5. The effect of reversing the amounts of reinforcement of two groups trained to asymptotic latencies on .05 and 2.4 gm. reinforcement. The extrapolation of the pre-change asymptotes as dotted lines provides a graphic picture of the extent of contrast effects, which take the form of areas between the dotted lines and the solid curves drawn by inspection through the experimental points. The extinction curves have also been drawn by inspection.

and 1/64, or 34 and 2 percent, respectively. The contrast produced by the fiftyfold change is significant, therefore, for the increment but not the decrement in amount of reinforcement.

Extinction after the change in amount.—The course of extinction after the change in amount of reinforcement is depicted in Figs. 2 and 5. The curves presented in Fig. 5 have been drawn by inspection through the experimental points, the variability being too great to warrant the fitting of equations. The experimental points in Fig. 2 have been connected by straight lines to help overcome the confusion of overlapping. A direct relation between asymptote level and resistance to extinction can be seen to exist in Fig. 2, the lower the asymptote the lower the latency tends to be for any extinction trial up to approximately the fourth trial, after which the three intermediate amount groups, 1.2, .60 and .20 gm., assume a common level of responding. The extreme amount groups, 2.4 and .05, remain separated throughout the ten trials of extinction, a possible manifestation of positive and negative contrast effects.

Application of the type of statistical analysis used previously brought out significant differences on at least the two-percent level between the extreme amounts, and between each extreme amount and at least one of the intermediate amounts. None of the differences between the three intermediate amounts was significant.

The extinction of the two groups subjected to the reversal of .05 and 2.4 gm. is shown in Fig. 5. Again, the strength of responding before extinction determines the strength during extinction. The effect of the eight days' reversal of amount is apparently greater than the 19 days training on the original amounts. Since all nine extinction trial values of the .05 group exceed those of the 2.4 group, the difference between them is significant at the one-percent level.

Theoretical.—If it is assumed that the log latency measure is inversely and linearly related to Hull's intervening variable ${}_{S}E_{R}$, and also, as a consequence of constant drive, inversely proportional to the quantitative construct habit strength (${}_{S}H_{R}$), then Hull's theoretical description of the growth of habit strength under different amounts of reinforcement becomes applicable to the empirical learning functions of this investigation. Hull (13) states that the law of simple habit strength acquisition is an exponential function of the number (N) of reinforcements, mathematically expressible as

$$ {}_{S}H_{R} = m(1 - e^{-iN}) $$

where i is a constant controlling the rate-of-approach of the learning curve to m, the limit of habit strength growth. Of the two constants, i and m, it is the latter which is postulated to be functionally related to the amount of reinforcement. Since Hull's acquisition equation and the equations fitted to the empirical learning curves of the present experiment are the the same general exponential form, a reasonably precise test of theory was possible. The latency asymptote (P_{∞}) corresponds to the limit of habit strength (m), and the constant C governing the learning curves' rate-of-approach to P_{∞} corresponds closely to Hull's constant i. The findings of this investigation with respect to these constants were completely in conformity with theoretical expectations. The limits of learning were found to be a positive monotonic function of the amount of reinforce-

ment, while the rate-of-approach to the limits remained constant.

Still further confirming evidence of Hull's amount-of-reinforcement hypothesis may be adduced from the data of another possible index of habit strength, resistance to extinction. This measure took the form of the extinction-equation constant, c, which represents the rate-of-approach to the final low level of responding. This constant also bore a positive monotonic relation to the amount of reinforcement, as theory predicted.

The systematization put forth in *Principles of behavior* (13) purports to account for the effects of a change in the amount of reinforcement at any time during the process of reinforcement. Hull provides an explicit mechanism for the increase in amount of reinforcement, but none for the corresponding case of amount-decrease, regarded as a more complex problem involving extinction effects. The mechanism proposed is that, at any trial after the start of acquisition on a given amount, the introduction of a greater amount will cause the curve of acquisition to rise as a simple growth function approaching the limit associated with the greater amount. The data of this experiment are in accord with this notion, although there is some indication that large increments introduce contrast effects requiring more elaborate explanation.

The experimental demonstration of the response strength decrement induced by lowering the quantity of reward provides a basis for speculation as to the kind of explanatory mechanism which would predict the effect and yet be compatible with other relationships within the theoretical structure. There are two obvious possibilities to be considered. The first is simply that further reinforcement under diminished amount reduces the limit of rise of habit strength to that associated with the lesser amount; and if it should be the case that this lower limit were already exceeded, ${}_{s}H_{R}$ would decrease exponentially with further reinforcements until that lower limit were reached.

The second possibility concerns a shut-off mechanism whereby ${}_{s}H_{R}$ remains constant when further reinforcements are with lesser amounts whose m values are not great enough to support the already existing ${}_{s}H_{R}$. The observed decrease of response strength accompanying reduction in reward is then accounted for in terms of the accumulation of reactive inhibition (I_R) and conditioned inhibition (${}_{s}I_{R}$), which builds up as a function of the work done in the situation.

Of these two possibilities, the second is the less likely for the following reasons. First, the radical spacing procedure eliminated the accumulation of reactive inhibition, (I_R), so that any response decrement would have to be attributed to conditioned inhibition (${}_{s}I_{R}$). But since ${}_{s}I_{R}$ is itself dependent upon accumulation of I_R, there is little likelihood of this factor being quantitatively strong enough to account for the demonstrated extinction effects. The second reason lies in the failure of this type of mechanism to account for the correlation between amount of reward and response decrement. There is no reason why the groups shifted to smaller amounts should have developed greater inhibitional tendencies to lengthen their latencies.

In summary, this means that the quantitative reduction in amount of reinforcement probably acts through the agency of a reduction in habit strength (${}_{s}H_{R}$), rather than through the build-up of inhibition (I_R and

${}_sI_R$), to depress the strength of response.

The most important consequences of this analysis become apparent when it is extended to apply to a special case of change-in-amount-of-reinforcement, that of reduction to zero amount—the operation of experimental extinction. It may be noted that in Hull's theoretical system extinction is attributed not to the draining off of ${}_sH_R$ by non-reinforced elicitations, but to the accumulation of the two kinds of inhibition. If the present analysis is correct, extinction it to be regarded both as a reduction in amount of reward involving a greater drainage of ${}_sH_R$ than less extreme reductions, and also as an accumulation of inhibition. While there are undoubtedly many differential consequences of these two alternative interpretations, a conclusive test of the 'drainage' view would consist of a demonstration of extinction with no evidence of inhibition present. In experimental terms this means eliminating I_R by reducing response work, spacing extinction trials at one-a-day, then testing for ${}_sI_R$ at various stages of extinction through the ability of ${}_sI_R$ to generalize to other responses (13, p. 267; 23) and its tendency to be reduced (disinhibited) by extraneous stimuli (13, p. 287; 14).

Up to this point, the complexities introduced by the evidence of the contrast effects have not received a theoretical treatment. Crespi (1, 2) has theorized extensively about the relationships between level of performance and amount of reinforcement, his conception being that the quantity of incentive acts upon behavior through variation of the drive (D) rather than habit strength (${}_sH_R$). Differences in the consequences of the two notions are not easily stated, since observable behavior is mediated by the product of these two intervening variables (13). Whether the operation of varying the amount of reinforcement more properly belongs with a class of deprivation operations (D) or reinforcement operations (N) cannot be determined at present by weight of evidence.

It is to be admitted that the primary principles of the Hullian system have some difficulty in incorporating the contrast effect, and furthermore, that Crespi's inferences of acquired, emotional drive states will have to be paralleled or adopted within Hull's system if it is to remain compatible with all existing relevant data. It has been shown in two investigations, now, with two different measures of response magnitude, that the operation of increasing and decreasing the quantity of reinforcement produces stronger and weaker responses than are attributable to primary reinforcement. The evaluation of any states of the organism constructed to account for these effects depends upon how well the further properties of these states can be worked out. Crespi's data and interpretation led him to assign short temporal courses to the states of 'elation' and 'depression,' although his curves show an empirical basis for this only with the negative contrast effects. The positive contrast effect and hence the state of 'elation,' in the present experiment and Crespi's, showed no signs of dissipating at the end of the eight trials under increased amount. The heightened extinction effects comprising the 'depression,' on the other hand, did not tend to last for the eight trials of diminished amount, for either latency or running speeds (1). The two acquired drive states, or perhaps the secondary reinforcing stimulus properties, whatever they might be,

must be such as to predict the differences apparent in temporal course.

Still another, and final, comparison to be made concerns the magnitude of the small-amount 'depression' effect, which Crespi found to be great enough to weaken behavior more than reduction to zero amount. The same observation did not hold for the conditions and measures of this experiment. Extinction training promptly elevated the latencies of animals trained under every condition of amount. A satisfactory theoretical account of the contrast effect remains to be given.

Summary and Conclusions

1. For the purpose of determining the effects of the parameter of amount of reinforcement on empirical acquisition and extinction functions involving a latency measure of response magnitude, seven groups of white rats were given well-spaced training on a simple elevated runway with amounts of food reinforcement weighing .05, .20, .40, .60, .80, 1.60, and 2.40 gm. When response latencies had reached stable asymptotes, the four groups receiving .20, .40, .80, and 1.60 gm. were extinguished with massed trials. The groups receiving the largest and smallest amount had their rewards interchanged, and were then given eight additional daily trials before extinction. Following acquisition, the 40 animals of the .60-gm. group were divided into five equated subgroups and given eight more daily trials with the following amounts: .05, .20, .60, 1.20, and 2.40 gm. These animals were also subjected to massed extinction trials.

2. The family of comparable learning curves generated by the varying amounts of reinforcement differed significantly with respect to equation constants representing the final limit or asymptote of latencies. The parameter representing the rate-of-approach to these limits remained unchanged.

3. A plot of log latency asymptote as a function of log grams of food yields a curve negative in slope and slightly positively accelerated over the range tested.

4. The greater resistance to extinction of the larger amount-groups was indicated by the systematic variation of the extinction equation constants controlling the rate-of-approach to the final level.

5. Changes in the amount of reinforcement of the intermediate amount group at the asymptote stage of acquisition were accompanied by corresponding changes in log latency levels such that a plot of the new levels against log grams of food yielded a curve similar to the simple amount function described above. Greater resistance to extinction was again found to be positively related to the amount of reinforcement on the previous acquisition trials.

6. The reversal of amounts of the .05- and 2.4-gm. groups produced a reversal in the latency levels, but a comparison of the pre-change and post-change asymptotes for the .05-gm. reinforcement showed the post-change to be significantly lower. This defines what has been called a positive contrast effect. A similar comparison for the 2.4-gm. reinforcement shows the post-change latencies to be higher, but not for a sufficient number of trials to be statistically significant. This tendency has been termed the negative contrast effect.

7. Certain theoretical implications and relations to previous findings are considered. It is suggested that quantitative reduction in amount of reinforcement reduces strength of response by reducing habit strength

($_sH_R$). The theory is advanced that extinction is to be regarded as a reduction in amount of reinforcement, and also, contrary to Hull's view, as involving a reduction in habit strength.

(Manuscript received August 16, 1948)

References

1. Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55**, 467–517.
2. Crespi, L. P. Amount of reinforcement and level of performance. *Psychol. Rev.*, 1944, **51**, 341–357.
3. Felsinger, J. M., Gladstone, A. I., Yamaguchi, H. G., & Hull, C. L. Reaction latency ($_sI_R$) as a function of the number of reinforcements (N). *J. exp. Psychol.*, 1947, **37**, 214–228.
4. Finger, F. W. The effect of varying conditions of reinforcement upon a simple running response. *J. exp. Psychol.*, 1942, **30**, 53–68.
5. Fitts, P. M. The effect of a large and a small reward as indicated by the resistance-to-extinction curve for the rat. *Psychol. Bull.*, 1940, **37**, 429–430.
6. Fletcher, F. M. Effects of quantitative variation of food-incentive on the performance of physical work by chimpanzees. *Comp. Psychol. Monogr.*, 1940, **16**, No. 82.
7. Gagné, R. M. The effect of spacing of trials on the acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1941, **29**, 201–216.
8. Gagné, R. M. External inhibition and disinhibition in a conditioned operant response. *J. exp. Psychol.*, 1941, **29**, 104–116.
9. Graham, C. H., & Gagné, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
10. Grant, D. A. New statistical criteria for learning and problem solution in experiments involving repeated trials. *Psychol. Bull.*, 1946, **43**, 272–282.
11. Grindley, G. C. Experiments on the influence of the amount of reward on learning in young chickens. *Brit. J. Psychol.*, 1929, **30**, 173–180.
12. Hull, C. L., Felsinger, J. M., Gladstone, A. I., & Yamaguchi, H. G. A proposed quantification of habit strength. *Psychol. Rev.*, 1947, **54**, 237–254.
13. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century Co., 1943.
14. Hunter, W. S. The disinhibition of experimental extinction in the white rat. *Science*, 1935, **81**, 77–78.
15. Kimble, G. A. Private communication.
16. Lawrence, D. H., & Miller, N. E. A positive relationship between reinforcement and resistance to extinction produced by removing a source of confusion from a technique that had produced opposite results. *J. exp. Psychol.*, 1947, **37**, 494–509.
17. Mote, F. A. Jr., & Finger, F. W. Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. *J. exp. Psychol.*, 1942, **31**, 57–68.
18. Schlosberg, H., & Solomon, R. L. Latency of response in a choice discrimination. *J. exp. Psychol.*, 1943, **33**, 22–39.
19. Solomon, R. L. Latency of response as a measure of learning in a "single-door" discrimination. *Amer. J. Psychol.*, 1943, **56**, 422–432.
20. Skinner, B. F. *The behavior of organisms: an experimental analysis.* New York: Appleton-Century, 1938.
21. Verplank, W. S. The development of discrimination in a simple locomotor habit. *J. exp. Psychol.*, 1942, **31**, 441–464.
22. Wolfe, J. B., & Kaplon, M. D. Effect of amount of reward and consummative activity on learning in chickens. *J. comp. Psychol.*, 1941, **31**, 353–361.
23. Youtz, R. E. P. The weakening of one Thorndikean response following the extinction of another. *J. exp. Psychol.*, 1939, **24**, 294–304.

FACTORS INFLUENCING THE OCCURRENCE OF REMINISCENCE: ATTEMPTED FORMAL REHEARSAL DURING THE INTERPOLATED PERIOD [1]

BY JOHN H. ROHRER [2]

University of Oklahoma

Introduction

Theories which attempt to account for reminiscence may be conveniently regarded as falling into two categories: first, that group of theories which attempts to explain the improved recall performance (i.e., the phenomenon of reminiscence) in terms of a shift, during the interpolated period, in the strength of the associative tendency existing between the S-complex and the R. Such changes in associative strength have been variously attributed to: (a) a maturation process (16); (b) informal rehearsal during the rest interval (2, 17, 18); and (c) a differential loss of associative strength for right and wrong responses (5).

A second class of theories attempts to explain reminiscence in terms of a temporally labile inhibitory factor which, during acquisition, interferes with performance. Reminiscence is said to be due to a removal, during the interpolated interval, of the inhibitory state thereby permitting the maximum effective associative tendency of the R to manifest itself during recall. Specific examples of this type of reminiscence theory are McClelland's (10) performance decrement theory, the physical fatigue theory suggested by Dashiell (3), and Hull's reaction inhibition theory (8). The study reported herein was designed: (1) to test the hypothesis that the phenomenon of reminiscence results from a building up of the S-R associative tendencies by rehearsal occurring during the period interpolated between acquisition and recall; and (2) to test further the hypothesis that the phenomenon of reminiscence results from the dissipation, during the interpolated period, of an inhibitory tendency which originally was built up during acquisition through frustration or non-reward.

Method

The 24 naive general psychology students who served as *Ss* in this experiment were required to learn three discrete lists of 14 nonsense syllables each to a criterion of 9 out of 14 correct responses. The lists were presented serially on a Missouri-type memory drum. During both the acquisition and recall sessions the *Ss* learned by the anticipation method. Under all conditions and at all sessions the syllables were exposed for two sec. each, with a two-sec. period between trials. The influence on recall of three distinct activities interpolated during five-min. intervals between the criterion learning trial and the recall trial was studied. Table I outlines the counter-

TABLE I

Learning Sequence for Experimental Groups

Learning Session	Exp. Cond. 1 List I	Exp. Cond. 2 List II	Exp. Cond. 3 List III
1st	A. $N = 8$	B. $N = 8$	C. $N = 8$
2nd	C. $N = 8$	A. $N = 8$	B. $N = 8$
3rd	B. $N = 8$	C. $N = 8$	A. $N = 8$
Total N	$N = 24$	$N = 24$	$N = 24$

[1] A summary of this study (13) was presented at the 1946 meeting of A.P.A.

[2] Purchase of the apparatus used in this study was made possible by a financial grant from the Faculty Research Committee of the University of Oklahoma. The writer wishes to express thanks to Ruth F. Logan who ran the *O*'s and aided in analyzing the data.

TABLE II

ACTIVITIES OF SUBJECTS UNDER EACH CONDITION

	Experimental Condition 1	Experimental Condition 2	Experimental Condition 3
Acquisition Activity	Learn List I	Learn List II	Learn List III
Interpolated Activity	Symbolic activity unrelated to acquisition activities for 5 min.	Read jokes for 5 min.	Formally rehearsed incompletely learned list for 5 min.
Recall Activity	One trial of recall on List I	One trial of recall on List II	One trial of recall on List III

balancing design which was used. It will be noted that such a design confounds practice effects and individual differences in the learning ability of the *Ss*. It does not provide for an estimate of the amount of score variability due to the intrinsic characteristics of the lists themselves. The lists were selected from Hull's list (7) in such a manner that the positional items in the three lists were matched as to associational difficulty. It would of course follow that each list had equal associational difficulty. The *Ss* were assigned to the A, B, and C groups through the use of a table of random numbers (9).

Table II outlines the experimental procedure followed. It will be noted that the influence on recall of three distinct interpolated activities was studied. The activities were interpolated during five-min. intervals between the criterion learning trial and the recall trials. Under experimental condition 1, the *Ss* learned List I to criterion, talked for five min. with a neutral interviewer about any topic *other* than his rote learning experiences, and was then given one additional recall trial on List I. Under experimental condition 2, the *Ss* learned List II to criterion, read jokes for five min., and then were given one additional recall trial on List II. This condition was included to provide a basis for comparison with traditional reminiscence studies. Under experimental condition 3, the *Ss* learned List III to criterion, attempted formal rehearsal of the learning material for the five-min. interpolated period, and then were given one additional recall trial on List III.

In an earlier experiment (14) it was demonstrated that the occurrence of reminiscence could be regulated by the degree to which a 'nondirective' or 'subject' or 'client-orientated' counseling technique was used to direct the flow of the *S*'s talk during the interpolated period. If the symbolic activity of the *S* centered around frustrated experiences encountered during the acquisition session, there was a decrement in performance during recall; if the symbolic activity of the *S* was unrelated to the learning or other frustrating experiences of the experimental session, reminiscence occurred. Experimental condition 1 of the present experiment was designed, in light of that earlier experiment, to provide a maximum reminiscence effect. During the interpolated interview the flow of talk of the *Ss* was directed or *structured* by the neutral counselor's[3] opening statement plus subsequent restatements; i.e., the interview was *structured* so that the *S* did not talk about his acquisition period experiences during the five-min. period. The procedure for the formal rehearsal activity interpolated under condition 3 was as follows: The *Ss* sat before the exposure apparatus on which 14 blank lines were exposed at two-sec. intervals with a two-sec. interval between trials. The lines corresponded in position to the syllables on the learning list. The beginning and end of the list were designated by an asterisk identical to the one used with the acquisition list. The *Ss* were instructed to at-

[3] The writer was the 'neutral' counselor who was summoned by buzzer from an adjoining room. He was summoned as soon as *S* reached the criterion score. The *S* was unaware that the counselor had been summoned. Approximately 45 sec. elapsed between the *S*'s final R, indicating he had reached the criterion, and the appearance of the counselor. The *E* was absent from the room during the interview. She returned to interrupt the interview approximately 30 sec. prior to the recall trial. The interviews were casual. The usual opening statement was an inquiry about the location of *S*'s home town followed by the minimum number of restatements required to keep the *S* explicitly verbalizing concerning experiences *other than* those encountered in the experiment or past experiences in which *S* had encountered failure. An analysis of the protocols of the interviews indicated that none of the *Ss* talked of identifiable failure experiences during the interpolated period.

tempt a subvocal rehearsal of the list they had just previously been learning. If rehearsal of a partially learned list is the variable responsible for the production of reminiscence, then the recall results of the *Ss* under condition 3 should demonstrate a marked amount of reminiscence, since the physical S-conditions during interpolation closely approximated the conditions which existed during original learning. To further assure similarity between acquisition and interpolated activities, the *Ss* under this condition were instructed to rehearse the list at a speed identical to the acquisition rate of presentation, thereby creating a condition that more nearly approximated the uniform practice interval existing during acquisition than would be possible under informal rehearsal circumstances. Empirically it is a well-established generalization that, other things equal, the more similar the S-complex during recall is to the S-complex present during learning, the better the resulting recall performance. Consequently, according to the informal rehearsal hypothesis, one would predict that the *Ss* under condition 3 would show a large amount of reminiscence, because of the physical similarity of the acquisition and interpolated activities to the recall activity.

Results

Table III presents a summary of the results obtained under the various experimental conditions. The acquisition performance of the three groups was essentially the same as evidenced by the *M*'s and *SD*'s for the various conditions. None of the observed differences between the acquisition means approached the levels of confidence commonly required for statistical significance. Column 7, showing the percent reminiscence, indicates that experimental group I was the only group which demonstrated an over-all increased score during recall. Twenty of the 24 *Ss* demonstrated an improved recall score under condition 1, 12 of them demonstrated an improved recall score under condition 2, and none of them demonstrated an improved recall score under condition 3.

It will be observed in Table III that there is a mean difference of +0.65 syllables between the mean criterion and recall scores for group I and a mean difference of −1.58 syllables between the mean criterion and recall scores for group III. A question may be raised with respect to the statistical significance of those observed mean differences. Student's *t*-test (9) for determining the significance of mean differences was applied to the data. The observed difference between the means of group I was found to be significant at the five percent level of confidence (*d.f.* = 46). The observed difference for group II was not statistically significant. The observed difference between the mean criterion and recall scores of group III was found to be significant beyond the one percent level of confidence. On the basis of

TABLE III

Summary of the Mean Recall Results Obtained Under the Various Experimental Conditions

Experimental Group	Type of Interpolated Activity During Five-min. Period	Mean Score on Criterion Trial of Training	*SD* Training	Mean Score on Recall	*SD* Recall	Mean Recall Score as % of Mean Criterion Score
I	Counseled away from situation	9.54	0.6	10.19	1.29	106%
II	Read jokes	9.45	0.7	9.26	1.66	96%
III	Rehearsed learned material	9.60	0.6	8.02	1.14	83%

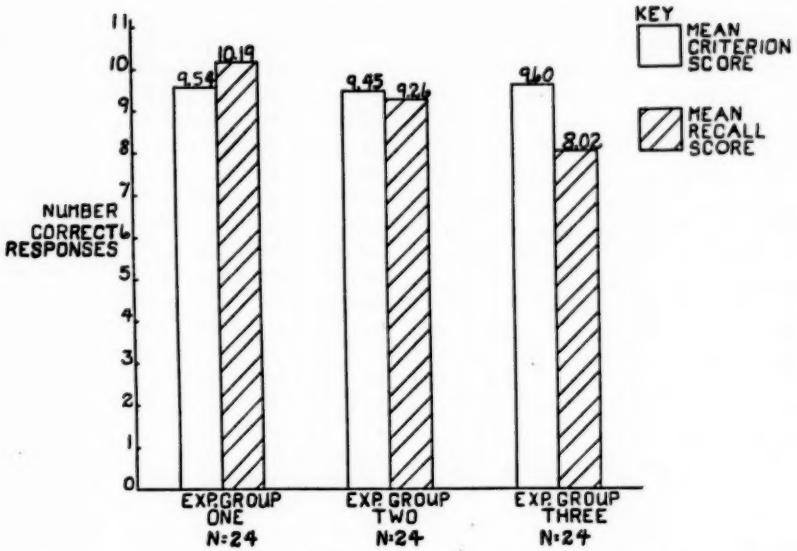

FIG. 1. A comparison of the mean criterion and recall scores obtained by the experimental groups

these results we may reject the hypothesis that the increased recall score obtained by group I and the decreased recall scores obtained by group III can be attributed to random errors. One may conclude that these observed differences were due to the independent variable manipulated during the interpolated interval, i.e., that the counseling activity facilitated recall and the formal rehearsal interfered with recall.

Fig. 1 presents the mean difference between the criterion and recall scores for each group. Again, it will be seen that the acquisition performance of all three groups was essentially the same, and that the only group which demonstrated an over-all improvement on recall was group I.

Table IV presents the results of an analysis of the mean recall differences obtained under the three experimental conditions. It will be noted that the difference between conditions I and II, in favor of I, was significant at between the 20 and 10 percent levels of confidence; that the difference between conditions II and III was significant at between the 20 and 10

TABLE IV

A STATISTICAL COMPARISON OF DIFFERENCES BETWEEN THE MEAN NUMBER OF CORRECT ANTICIPATIONS MADE DURING RECALL

Comparison of Experimental Conditions	Difference in Favor of Condition	*d.f.*	*t*	Level of Confidence at Which Significant
1 and 2	1	46	1.5130	Between 20 and 10%
1 and 3	1	46	2.9045	1%
2 and 3	2	46	1.4893	Between 20 and 10%

percent levels of confidence; and that the difference between conditions I and III was significant at the one percent level of confidence.

Discussion of Results

In view of the data obtained in this experiment the hypothesis is highly untenable that reminiscence can be attributed to rehearsal occurring during the interpolated period. The rehearsal activity in which the *S*s engaged under condition 3 resulted in a recall decrement, rather than in the increment which would be predicted from the rehearsal theory. Conversely, the activity interpolated in condition 1 would minimize the opportunity for rehearsal; yet it was under this condition that the greatest amount of reminiscence was obtained. This latter finding is in general agreement with a study by McGeoch (12), who found that informal rehearsal did not increase the amount of observed reminiscence. The data do favor the hypothesis that the phenomenon of reminiscence is due to the dissipation of an inhibitory performance factor which, during acquisition, interfered with the elicitation of the learned R. According to this latter interpretation, the increased score obtained during recall resulted from the dissipation, during the interpolated period, of a temporally labile inhibitory factor. An attempt to define further the nature of this inhibitory factor responsible for reminiscence leads to conflicting points of view.

Hull's (8) reactive inhibition theory assumes that an inhibitory tendency, I_R, is built up during acquisition *with each R made, regardless of whether or not the R is correct.* Further, it is assumed that this inhibitory tendency reduces the effective excitatory tendency of the R; moreover, it is assumed that the strength of the inhibitory tendency weakens continuously in time. Hence, from massing the acquisition trials (a necessary condition for the production of reminiscence) there results a greater accumulation of the inhibitory tendency, which, in turn, reduces the effective excitatory tendency as evidenced by a decrease in the observed strength of the R. The interpolated interval, in turn, provides the opportunity for the dissipation of the inhibitory state, thereby permitting the manifestation at recall of the complete uninhibited effective excitatory tendency of the acquired R. It should follow from this theoretical formulation that the amount of reminiscence exhibited should be directly related to the number of responses made during acquisition. An analysis of the data under condition 1 of this experiment indicated that the number of responses made during acquisition was not related to the amount of reminiscence demonstrated. The coefficient of correlation (rho) between the total number of responses made during acquisition and the amount of reminiscence exhibited during recall was $-.14$. This result argues against Hull's reactive inhibition theory.

A second possible variable which could be responsible for a temporally labile inhibitory state is the *number of frustrations or non-rewards* that the *S* experiences during acquisition. It is to be noted in this connection that frustration, as that concept is operationally defined by Dollard *et al.* (4), could occur during acquisition in at least two different places: first of all, non-reward would result when the *S* made an anticipation of a particular syllable which later was shown to be incorrect; secondly, non-reward could be a variable introduced as a result of a set, on the part of the *S*, to *learn the list.* In fact, to learn the list is

the over-all *goal* for which the *S* was set by the experimental directions. To the extent that this goal was not achieved, frustration had occurred. The results obtained under experimental condition 3 of the present experiment would specifically point to an interpretation in terms of frustration, for under this condition the *Ss* continued to be frustrated during the five-min. interpolated period, since they were attempting to recall material, a part of which was below the limen of recallability. It will be seen in Table III that it was under this condition that the poorest recall score was obtained.

The results obtained under condition 1 would also support the hypothesis that frustration was the variable responsible for the development of the inhibitory tendency, since the interpolated activity under this condition was of such a nature as to minimize the amount of direct or symbolic frustration experienced. It was under this condition that the greatest amount of reminiscence was obtained. Finally, under condition 2 the frustration variable was not controlled during the interpolated period and it was under this condition that the greatest recall variability was observed. One-half of the *Ss* under condition 2 did exhibit higher scores at recall. An hypothesis which would account for the greater variability observed under this condition is as follows: those 12 *Ss* who exhibited reminiscence actually engaged in the instructed interpolated activity of joke reading, which would minimize the amount of symbolic frustration, permit maximum dissipation of the inhibitory state, and, hence, favor reminiscence as demonstrated in the results obtained under condition 1. Conversely, the 12 *Ss* under condition 2 who showed no improvement or a decrement in recall score failed to engage in the instructed activity of joke reading, but instead looked at the book and engaged in symbolic rehearsal of the original learning list thereby experiencing continued symbolic frustration. As demonstrated, in the results obtained under condition 3, such symbolic activities when resulting in non-reward, will result in a recall decrement.

The assumption that frustration is the variable responsible for the development during acquisition of an inhibitory state is further suggested in the results of an earlier study (14), in which it was demonstrated that *Ss* who, during an interpolated period, engaged in verbal activities related to the unsuccessful acquisition of a partially learned list showed a decrement in subsequent recall, while *Ss* who, during an interpolated period, engaged in verbal activities totally unrelated to the acquisition of a partially learned list did exhibit reminiscence on subsequent recall. The explanation offered of those data was to the effect that the verbal responses centering around the 'unsuccessful' activity were in themselves frustrating and hence resulted in a further accumulation of the inhibitory state built up during acquisition, which was reflected in the lowered scores on recall.

McClelland (11) has reported an experiment in which the *Ss* who showed the greatest amount of reminiscence were *Ss* who, during acquisition, made the greatest number of total correct anticipations. This result was verified under condition 1 of the present study. At first glance, such a result would seem to be damaging to a point of view which accounts for reminiscence in terms of an inhibitory factor resulting from frustration occurring during the acquisition series.

However, when one examines the conditions responsible for the production of reminiscence, it becomes obvious that for reminiscence to occur at all, the associative strength existing between the S-complex and the R must be sufficiently great for the response to be above the recall limen at the time that recall is attempted. It is to be noted that the acquisition criterion which is usually used (and was used in the present study) *leaves uncontrolled both the total number of rewards and frustrations which may occur during an acquisition series.* It has been adequately demonstrated (**8**) that one of the important experimental variables determining the strength of an association is the number of rewards or reinforcements. Hence, it would be expected that once the inhibitory factor had been dissipated during the interpolated period, those S-R associations with greater associative strength would more probably manifest themselves at recall and thereby produce the phenomenon of reminiscence. It is to be noted that reminiscence is conceived of as being a function of both the excitatory and inhibitory tendencies, according to the frustration hypothesis presented above.

Hovland has reported data (**6**) that demonstrate that under conditions of massed learning trials the number of error (i.e., *frustrated*) responses made increases greatly as compared with the number observed under distributed practice conditions. The massing of learning trials is an essential and necessary condition for the production of the phenomenon of reminiscence.

One further point should be noted concerning the attempt to explain reminiscence in terms of a rehearsal theory. In situations for which the Ballard (**1**) and Williams (**17**) experimental design is a prototype, one must answer the difficult question of how the *S*s can rehearse, during the interpolated period, that part of the list which they cannot recall at the start of the interpolated period. It is that unrecalled portion of the list that must be brought above the limen of recallability before reminiscence can occur.

Summary

This experiment was designed (1) to determine more specifically the relationship between attempted rehearsal of a partially learned list during an interpolated period and the subsequent recall of that list; and (2) to test further the hypothesis that the phenomenon of reminiscence results from the dissipation, during the interpolated period, of an inhibitory tendency which originally was built up during acquisition through frustration or non-reward.

Data are presented which show that attempted rehearsal of *a partially learned list* of verbal materials during an interpolated period will result in a decrement in subsequent recall performance. The data are interpreted as supporting the hypothesis that the phenomenon of reminiscence results from the dissipation of an inhibitory tendency built up during acquisition as a result of non-reward or frustration.

(Manuscript received September 6, 1948)

References

1. Ballard, P. B. Obliviscence and reminiscence. *Brit. J. Psychol. Monogr. Suppl.*, 1913, **1**, No. 2.
2. Buxton, C. E. The status of research in reminiscence. *Psychol. Bull.*, 1943, **40**, 313–337.
3. Dashiell, J. F. *Fundamentals of objective psychology.* New York: Houghton Mifflin, 1928.
4. Dollard, J., Doob, L. W., Miller, N. E.,

MOWRER, O. H., & SEARS, R. R. *Frustration and aggression.* New Haven: Yale Univ. Press, 1939.

5. HOVLAND, C. I. Experimental studies in rote-learning theory. VI. Comparison of retention following learning to same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, **26**, 568–587.
6. HOVLAND, C. I. Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and by distributed practice. *J. exp. Psychol.*, 1938, **23**, 172–190.
7. HULL, C. L. The meaningfulness of 320 selected nonsense syllables. *Amer. J. Psychol.*, 1933, **45**, 730–734.
8. HULL, C. L. *Principles of behavior.* New York: D. Appleton Century Co., 1943.
9. LINDQUIST, E. F. *Statistical analysis in educational research.* Boston: Houghton Mifflin, 1940.
10. MCCLELLAND, D. C. Studies in serial verbal discrimination learning. I. Reminiscence with two speeds of pair presentation. *J. exp. Psychol.*, 1942, **31**, 44–56.
11. MCCLELLAND, D. C. Studies in serial verbal discrimination learning. II. Retention of responses to right and wrong words in a transfer situation. *J. exp. Psychol.*, 1942, **31**, 149–162.
12. MCGEOCH, G. O. The conditions of reminiscence. *Amer. J. Psychol.*, 1935, **47**, 65–89.
13. ROHRER, J. H. Factors influencing the occurrence of reminiscence: Attempted formal rehearsal during the interpolated period. *Amer. Psychologist*, 1946, **1**, 259.
14. ROHRER, J. H., & RAYHILL, C. A. Some relationships between reminiscence and catharsis. *Amer. Psychologist*, 1946, **1**, 460.
15. WARD, L. B. Reminiscence and rote-learning. *Psychol. Monogr.*, 1937, **49**, No. 220.
16. WHEELER, R. H. *The science of psychology: an introductory study.* New York: Crowell, 1929.
17. WILLIAMS, O. A. A study of the phenomenon of reminiscence. *J. exp. Psychol.*, 1926, **9**, 368–387.
18. WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.

REMINISCENCE IN PURSUIT-ROTOR LEARNING AS A FUNCTION OF LENGTH OF REST AND OF AMOUNT OF PRE-REST PRACTICE

BY ARTHUR L. IRION

University of Illinois[1]

Introduction

Reminiscence may be defined as that increase in the proficiency of performance of a partially learned act which is attributable to the effects of an interpolated rest period. Following the studies of Ward (**10**) and Hovland (**5, 6, 7**), this definition has meant, in terms of the experimental paradigm reproduced below, that the post-rest proficiency of Group B is greater than the proficiency of Group A for the equivalent trials.

Group A
Original Learning
No Rest
Relearning

Group B
Original Learning
Rest
Relearning

In the pursuit-rotor situation, reminiscence has been shown to be some inverse function of the amount of pre-rest practice by Bell (**3**). Based upon somewhat different experimental conditions, Ammons' (**2**) results indicate that, as amount of pre-rest practice increases, amount of reminiscence at first increases and then decreases. In a related study, Buxton (**4**) demonstrated that amount of reminiscence decreases as the degree of pre-rest proficiency increases.

In studying reminiscence as a function of the length of the rest interval, Bell (**3**) varied the length of a single rest period introduced after either 5 or 15 one-min. trials. Five lengths of rest interval were used, the lengths ranging between 10 min. and 30 hours. After five trials of original practice, the various lengths of rest were followed by approximately equal initial increases in score. The longer rests, however, caused effects which persisted for a longer time during relearning. Following 15 trials of original learning, short rests appeared to be beneficial, but longer rests caused an initial post-rest decrement which was followed by a quick recovery during relearning.

Melton (**9**) introduced a single rest period of either 20 min., two days, or two weeks early in massed practice on the Koerth pursuit-rotor. Under these conditions, the 20-min. rest yielded the greatest amount of reminiscence early in relearning, but the two-day rest yielded the best final performance. These results are in general agreement with the findings of Bell (**3**), both with respect to the optimum length of rest intervals introduced early in practice and also with respect to his findings relative to the relationship between the length of rest and the slope of the relearning function.

Ammons (**2**) has obtained results which indicate that amount of reminiscence is a negatively accelerated increasing function of the length of the rest period. Amount of reminiscence was found to increase rapidly as the length of rest was increased up to a rest interval of approximately five min. in duration.

The warming-up effect, as it applies

[1] The research on which this paper is based was performed in the Laboratories of Psychology at the University of Missouri. The writer acknowledges his indebtedness to Dr. Kenneth W. Spence and to Dr. Willard R. Thurlow. The experimental results of this study were contained in a dissertation presented by the writer in partial fulfillment of the requirements for the Ph.D. degree, State University of Iowa, 1947.

to retention, may be defined in terms of the steeper slope of the relearning curve as compared with the slope of the original learning curve at the same level of initial proficiency. In cases where relearning performance is superior to original learning performance, as in the case where reminiscence occurs, the comparison is made between the slope of the relearning curve and the slope of the learning curve for a no-rest control group at a corresponding level of proficiency. Bell (3) noted that amount of warming-up effect, as measured by the difference in score between the first and second relearning trials, is some function of the length of the rest interval. His results suggest that, as length of the rest interval increases, the amount of warming-up effect tends at first to increase and then to decrease.

In his theoretical analysis of the motor learning situation, Ammons (1) speaks of amount of warm-up decrement (that is, amount of loss of proficiency attributable to the need for warming-up) rather than of the amount of the warming-up effect. In this analysis, Ammons assumes that amount of warm-up decrement is a negatively accelerated increasing function of the length of the retention interval. Ammons has further theorized that the warming-up factor may be identified with a loss of set to perform the activity involved. In other words, as the length of the retention interval increases, the various postural and attentive adjustments necessary for optimum performance tend to be lost. The warming-up period is identified with the time during relearning in which these adjustments are regained. According to this analysis, the greater the amount of pre-rest practice, the greater will be the number of these adjustments which will be acquired. Since loss of these adjustments is assumed to be a simple decay function, it follows that amount of warming-up effect should be greater following large amounts than following small amounts of original practice. Data obtained by Ammons (2) lend confirmation to this implication.

The present study is designed to extend the previous findings with respect to reminiscence as a function of amount of pre-rest practice and of the length of the rest period, and to yield some additional information regarding the warming-up effect and the conditions under which it occurs.

Apparatus and Procedures

Apparatus.—The apparatus consisted of a standard Koerth pursuit-rotor revolving at 60 rpm, an electric trial-timing device, and a Standard electric clock. The apparatus was so arranged that trials of a standard length with a constant length of intervening rest could be presented automatically. The basic time intervals employed throughout the experiments consisted of 25-sec. trials separated by five-sec. rests. All practice was given under these work-rest conditions. Two sec. before the beginning of each trial, a signal light mounted on the apparatus flashed as a warning to the *S*.

Scoring was accomplished by means of a Standard electric clock which was wired into circuit through the stylus, target, and pursuit-rotor relay. When the turntable was being turned, the clock would run whenever the stylus was in contact with the target. As soon as the pursuit-rotor motor was turned off at the end of a trial, the circuit to the clock was broken. Scores were recorded to the nearest 0.01 sec., the finest calibration on the clock. Unless otherwise specified, all scores in these experiments may be interpreted as the length of time in sec. and hundredths of sec. that the *S* kept the stylus in contact with the target during one 25-sec. trial.

Illumination was constant for every *S* in the experiment and consisted of the light provided by a 250-watt electric bulb suspended 4.5 ft. above the center of the pursuit-rotor turntable.

Subjects.—All of the *S*s who participated in these experiments were undergraduate students at the University of Missouri who volunteered their participation from the courses in general psychology. None of the *S*s had had previous experience with the Koerth pursuit-rotor.[3] No pre-training or trial runs were given to any *S*.

[3] Two peculiarities of the sample should be noted. The requirement that all *S*s be naive with respect to the pursuit-rotor situation excluded from the experiment a number of former Army Aviation Cadets who had received practice on this apparatus as a portion of aircrew classification testing. Secondly, it should be noted that women as well as men served in the

During preliminary investigational work, one *S* was discovered who was unable to learn the pursuit-rotor task. In order to avoid the inclusion of such *S*s in the experiments, a criterion was established for the selection of *S*s. Any *S* was discarded from the experiment if, after 10 trials, he had failed to attain a score of 1.00 sec. on any single trial and provided that he was not, by the end of the 10th trial, making the orthodox circular following motions.

One hundred twenty *S*s were used in the experiments, 15 being assigned to each of eight experimental conditions.

Rest Interval Activity.—During the rest intervals the *S*s were permitted either to read or to talk with the experimenter. They were not, however, permitted to talk about the pursuit-rotor task, nor were they permitted to inspect or manipulate the apparatus in any way. All questions concerning the nature of the experiment were referred to the end of the experimental session.

TABLE I

OUTLINE OF THE CONDITIONS OF EXPERIMENT I

Condition	Number of Pre-Rest Trials	Length of Rest	Number of Relearning Trials	*N*
1	10	5 min.	5	15
2	20	5 min.	5	15
3	30	5 min.	5	15
4	40	5 min.	5	15
5	40	No Rest	5	15

The Experimental Conditions.—This research is divided into two semi-independent experiments. The conditions of Experiment I, reminiscence as a function of the amount of pre-rest practice, are outlined in Table I. The 75 *S*s who served in this experiment were assigned at random to the five experimental groups. Each group contained 15 *S*s. Three *S*s were discarded from the experiment due to an inability to learn, as determined by the application of the criterion previously noted.

TABLE II

OUTLINE OF THE CONDITIONS OF EXPERIMENT II

Condition	Number of Pre-Rest Trials	Length of Rest	Number of Relearning Trials	*N*
A	20	No Rest	5	45
B	20	0.5 min.	5	15
C	20	1.0 min.	5	15
D	20	3.0 min.	5	15
E	20	5.0 min.	5	15

The conditions of Experiment II, reminiscence as a function of the length of the rest interval, are outlined in Table II. One hundred five *S*s served in this experiment. Sixty of these *S*s (those in conditions A and E) had been run in connection with Experiment I, above. Consequently, random assignment of *S*s to all conditions was not possible. Random assignment was made as between conditions A and E, and random assignment was also made as between conditions B, C, and D. However, *S*s in conditions A and E had no chance of being assigned to conditions B, C, and D, and vice versa. One additional *S* was discarded from this experiment due to an inability to learn.

RESULTS

Experiment I.—The mean cumulative scores for the experimental and control groups for the five trials preceding and the five trials following the rest interval are given in Table III.[3] It will be noted that the differences presented in the extreme right hand column of this table (for the post-rest trials) may be considered as measures

experiments. This arose as a result of the preliminary testing of 10 men and 10 women, the results of which indicated approximately equal mean performance as between the sexes. Later analysis of results indicated the presence of a systematic sex difference. Since the sexes were assigned at random to the conditions of the experiment, no systematic error should have been introduced as a result of the inclusion of both sexes. However, this factor would serve to increase the variability of the results that were obtained, a condition which should be borne in mind when interpreting the results.

[3] The scores presented in Table III and Table IV were obtained by summing the sums of scores for five trials and dividing by the product of the number of *S*s and the number of trials. In Table III it will be noted that the *N* varies between the several control groups. This is because it was sometimes possible to employ the original learning scores of *S*s in other experimental groups as control scores. Thus, the control group for condition 1 is composed not only of the *S*s in condition 5 but also of *S*s in conditions 2, 3, and 4.

TABLE III

A Comparison of Experimental and Control Group Scores Before and After the Introduction of the Five-Minute Rest Interval.
(Mean Cumulative Score for Five Trials)

Five Trials Before Rest

Condition	Experimental Group Means	*N*	Control Group Means	*N*	Difference (In Original Learning)
1 (10 Trials)	3.22	15	2.89	60	+0.33
2 (20 Trials)	4.14	15	4.15	45	−0.01
3 (30 Trials)	5.98	15	6.10	30	−0.12
4 (40 Trials)	7.44	15	7.38	15	+0.06

Five Trials After Rest

Condition	Experimental Group Means	*N*	Control Group Means	*N*	Difference (Amount of Reminiscence)
1 (10 Trials)	8.55	15	3.39	60	+5.16
2 (20 Trials)	10.80	15	5.37	45	+5.43
3 (30 Trials)	12.17	15	7.11	30	+5.06
4 (40 Trials)	12.04	15	8.29	15	+3.75

of reminiscence. The differences between the experimental and control groups are relatively small for the five trials preceding the introduction of the rest interval (none of these differences approaches statistical significance), whereas the differences following the rest are relatively greater (all are significant at the one percent level of confidence as determined by the *t*-test of significance). The differences in the right hand column of Table III for the post-rest trials are presented graphically in Fig. 1.[4]

When each relearning trial was considered separately, significant amounts of reminiscence were obtained on all relearning trials under all conditions with the exception of the first relearning trial under conditions 3 and 4 and the fourth relearning trial under condition 4.

Experiment II.—The mean cumulative scores for the experimental and control groups for the five trials preceding and the five trials following the rest interval are given in Table IV. It will be noted that the differences presented in the extreme right hand column of Table IV (for the post-rest trials) may be considered as measures of amount of reminiscence. The differences between the experimental and control groups are relatively small for the five trials preceding the introduction of the rest interval (none of these differences approaches statistical significance), whereas the differences following the introduction of the rest interval are relatively greater. The differences in the right hand column of Table IV (for the post-rest trials) are presented graphically in Fig. 2. It may be seen that, for the lengths of rest used in this experiment, the

[4] It will be noted in Fig. 1 that an experimental point (labeled 'A') is shown for two pre-rest trials. This point was obtained from a group of seven *Ss* who were run under the same general experimental conditions as the other *Ss* in Experiment I. However, these *Ss* were not randomly assigned over the conditions of this experiment, and hence are not included in the general treatment of results.

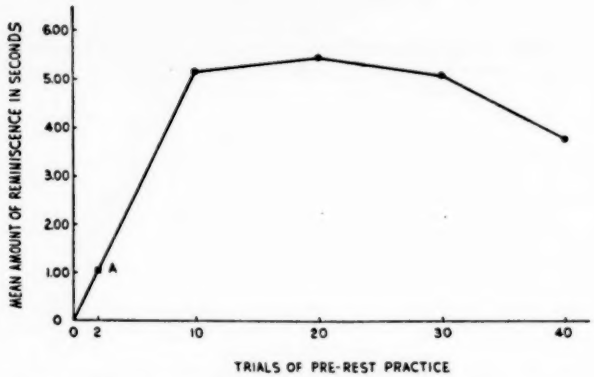

FIG. 1. Reminiscence as a function of amount of pre-rest practice

amount of reminiscence is an increasing, negatively accelerated, function of the length of the rest interval.

In Table V the differences between the experimental and control group means are given for each of the relearning trials. These data differ from those in Table IV in that they deal only with the trials following the introduction of the rest period, and in that the results are broken down by trials. The statistical significance of the differences presented in Table V was calculated by means of a simple analysis of variance for each of the relearning trials considered separately. Significance well beyond the one percent level of confidence was obtained in every case, thus establishing the significant presence of reminis-

TABLE IV

A COMPARISON OF EXPERIMENTAL AND CONTROL GROUP SCORES BEFORE AND FOLLOWING THE INTRODUCTION OF REST INTERVALS.
(Mean Cumulative Score for the Five Relearning Trials)

Five Trials Before Rest					
Condition	Experimental Group Means	*N*	Control Group Means	*N*	Difference (In Original Learning)
B (0.5 Min.)	3.94	15	4.15	45	−0.21
C (1.0 Min.)	4.16	15	4.15	45	+0.01
D (3.0 Min.)	4.79	15	4.15	45	+0.64
E (5.0 Min.)	4.14	15	4.15	45	−0.01

Five Trials After Rest					
Condition	Experimental Group Means	*N*	Control Group Means	*N*	Difference (Amount of Reminiscence)
B (0.5 Min.)	6.13	15	5.37	45	+0.76
C (1.0 Min.)	7.36	15	5.37	45	+1.99
D (3.0 Min.)	10.29	15	5.37	45	+4.92
E (5.0 Min.)	10.80	15	5.37	45	+5.43

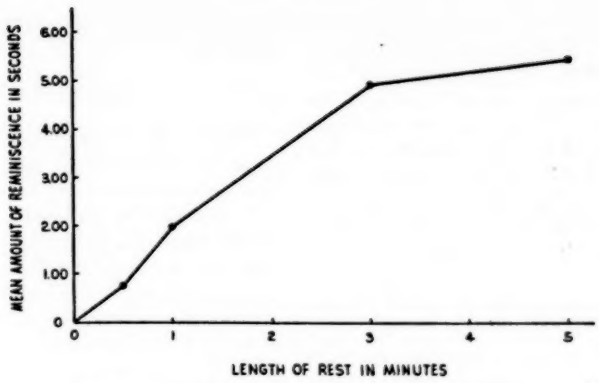

FIG. 2. Reminiscence as a function of length of rest

cence upon each of the relearning trials.

Warming-up Effect.—The warming-up effect may be defined in terms of the relatively steeper slope of the relearning curve as compared with the slope of the original learning curve at the same level of initial performance.

TABLE V

DIFFERENCES BETWEEN THE EXPERIMENTAL AND CONTROL GROUP MEANS (AMOUNT OF REMINISCENCE) FOR EACH RELEARNING TRIAL UNDER EACH OF THE CONDITIONS OF EXPERIMENT II

Condition	Relearning Trials					
	1	2	3	4	5	Mean
B (0.5 Min.)	0.58	0.76	1.14	0.61	0.68	0.75
C (1.0 Min.)	2.66	2.42	1.62	2.00	1.26	1.99
D (3.0 Min.)	4.16	5.37	5.32	5.19	4.52	4.91
E (5.0 Min.)	4.03	5.53	5.98	5.62	5.98	5.43

In the case of the experiments reported here, the acquisition curve was linear.[8] Because of this fact, differences in the slopes of the relearning curves, as between the conditions within the two experiments, may be interpreted as differences in amount of warming-up effect. Table VI shows the amount of change between the first and second relearning trials for the conditions of both experiments. It will be seen that amount of warming-up effect is, in general, an increasing function both of amount of pre-rest practice and of length of the rest interval.

Lindquist (8) has outlined a method for analyzing the significance of such slope differences as these. This

TABLE VI

AMOUNT OF CHANGE BETWEEN THE FIRST AND SECOND RELEARNING TRIALS FOR THE CONDITIONS OF EXPERIMENTS I AND II

Experiment	Condition	Mean Score on First Relearning Trial	Mean Score on Second Relearning Trial	Difference (Slope)
I	10 trials	7.19	8.99	+1.80
	20 trials	9.10	10.85	+1.75
	30 trials	9.77	12.61	+2.84
	40 trials	9.42	12.93	+3.51
II	No Rest	5.08	5.32	+0.24
	0.5 Min.	5.65	6.08	+0.43
	1.0 Min.	7.73	7.73	0.00
	3.0 Min.	9.23	10.69	+1.46
	5.0 Min.	9.10	10.85	+1.75

[8] The line of best fit, as determined by the method of least squares, was given by:

Score = 1.55 + 0.159 (No. Trials).

method tests the specific hypothesis that the differences in the two population means are the same (but not necessarily zero) for all values of the x variable. Taking the two extremes of the experimental variable in Experiment I (10 and 40 pre-rest trials) and considering all five of the relearning trials under both of these conditions, the hypothesis that is being tested is that, for the two populations involved, the slopes of the relearning curves are the same. This hypothesis is tested by means of the F-ratio, the conditions by trials interaction variance being used as the numerator and the combined subjects by trials interaction variance being used as the denominator. Should the hypothesis regarding the slopes of the relearning functions be true, this ratio should be distributed as F for the corresponding degrees of freedom, in this case, 4 and 112. Since computation of the ratio yields a value of 20.57, which is an extremely unusual value of F, the hypothesis concerning the slopes of the relearning curves may be rejected at a very high level of confidence.[6] An analysis of the slopes of the relearning curves for Experiment II involved the extreme experimental conditions A and E. The result of this analysis was an F-ratio of 27.67 for 4 and 232 degrees of freedom, which indicates that the slope of the relearning curve, and hence, the amount of warming-up effect, is a function of the length of the rest interval.

[6] It may be wondered why this analysis was performed only on two of the four experimental conditions. Interpretatively, however, no more conclusions could be drawn from the four condition analysis than from the two condition analysis. In either case, all that a significant value of F would demonstrate would be that the slope of the relearning curve is some function of the amount of pre-rest practice. It should be noted, however, that the selection of the two conditions for such an analysis must be made independently of the experimental results.

Conclusions

The results of Experiments I and II may be summarized under the following headings:

1. Significant amounts of reminiscence occurred following a five-min. rest which was introduced after 10, 20, 30, and 40 trials of practice on the Koerth pursuit-rotor under the conditions of Experiment I.

2. In considering the cumulative mean of the five relearning trials for each of the conditions of Experiment I, it may be concluded that, as amount of pre-rest practice increases, amount of reminiscence at first increases and then decreases.

3. Amount of warming-up effect increases as a function of the amount of pre-rest practice under the conditions of Experiment I.

4. Reminiscence was found to be a function of the length of the rest interval, the effects of rest being still measurable after five relearning trials.

5. Within the conditions of Experiment II, amount of reminiscence appears to be a negatively accelerated, increasing function of the length of the rest period, when the cumulative mean of the five relearning trials is used in the determination of the amount of reminiscence.

6. Amount of warming-up effect increases as some function of the length of the rest interval.

(Manuscript received September 2, 1948)

References

1. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.

2. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance with

continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–411.

3. Bell, H. M. Rest pauses in motor learning as related to Snoddy's hypothesis of mental growth. *Psychol. Monogr.*, 1942, **54**, No. 1.
4. Buxton, C. E. Level of mastery and reminiscence in pursuit learning. *J. exp. Psychol.*, 1943, **32**, 176–180.
5. Hovland, C. I. Experimental studies in rote-learning theory. I. Reminiscence following learning by massed and distributed practice. *J. exp. Psychol.*, 1938, **22**, 201–224.
6. Hovland, C. I. Experimental studies in rote-learning theory. II. Reminiscence with varying speeds of syllable presentation. *J. exp. Psychol.*, 1938, **22**, 338–353.
7. Hovland, C. I. Experimental studies in rote-learning theory. IV. Comparison of reminiscence in serial and paired associate learning. *J. exp. Psychol.*, 1939. **24**, 466–484.
8. Lindquist, E. F. Goodness of fit of trend curves and significance of trend differences. *Psychometrika*, 1947, **12**, 65–78,
9. Melton, A. W. The effect of rest pauses on the acquisition of the pursuitmeter habit. *Psychol. Bull.*, 1941, **38**, 719. (Abstract)
10. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 220.

PERFORMANCE AND REMINISCENCE IN MOTOR LEARNING AS A FUNCTION OF THE DEGREE OF DISTRIBUTION OF PRACTICE

BY GREGORY A. KIMBLE

Brown University

INTRODUCTION

In a previous paper, the author (5) presented a two-component theory of inhibition designed to handle certain phenomena of motor learning. This theory may be regarded as an extension of that suggested by Hull (3), since it makes use of the same major constructs, reactive inhibition (I_R) and conditioned inhibition (sI_R). An outline of the theory which is sufficient for the derivation of certain experimental predictions follows.

Reactive inhibition is regarded as a response-produced drive state for which the goal response is resting. It is a need to cease responding which results from or accompanies all behavior whether reinforced or not. Since I_R is a response-produced drive, it is not surprising to find that the important variables determining its strength at any moment are closely related to the characteristics of the response in the organism's immediate past. In general, we may say that the more *effortful* the behavior in which the organism has just been engaged, the greater the amount of I_R present, and consequently the greater the need to cease activity. This effortfulness of behavior in turn depends upon two independent variables, the physical energy required to perform the task at hand and the length of time the organism has been forced to work without rest. The first of these independent variables has received little experimental attention. In the case of the second, however, Ammons (2) and Kimble (5) have presented findings that indicate that reactive inhibition increases as a function of practice and that the course of the increase is negatively accelerated with an asymptotic value of I_R being reached after seven or eight min. of continuous practice.

Since I_R is a drive, it is postulated to have, in common with other drives, the following characteristics:

1. It will be reducible by the appropriate goal response, resting; that is, it will dissipate or decay during periods of rest. The experiment of Kimble and Horenstein (8) may be cited as a confirming test of this hypothesis as well as of Hull's suggestion that the course of this decay is exponential.

2. The decay of reactive inhibition is a reinforcement, since this decay is actually the reduction of a drive.

3. The accumulation of a certain (critical, threshold) amount of I_R will automatically produce resting. As with other drives, the development of a certain, critical drive strength will lead to the appropriate drive reducing behavior if the necessary goal object is available. In the case of reactive inhibition, the opportunity to rest is never absent from the work or learning situation. All that is necessary to produce the resting response is for a sufficient amount of inhibition to accumulate. The amount of I_R that is sufficient is postulated to depend upon the strength of other, opposing motives. In motor learning, the resting response will occur when the amount of reactive inhibition becomes

relatively greater than the strength of the motives leading to the continuation of practice.

From this hypothetical interaction between reactive inhibition and other motives in the learning experiment, an important deduction is the following: after the early practice trials in the motor learning situation, the amount of I_R will tend to remain constant at some maximal level. The accumulation of I_R above this level will be prevented by the occurrence of the resting response which reduces reactive inhibition. On the other hand, when I_R drops below this level, practice will resume and I_R will begin to develop again.

From the characteristics of I_R just proposed, the development of ${}_sI_R$ can be derived. Conditioned inhibition is postulated to be a habit. It is a resting habit which is acquired because resting behavior is reinforced by the reduction of I_R. The situation seems most appropriately analyzed in conditioning terms. The drive, reactive inhibition either produces, is accompanied by or is identifiable with drive stimuli which are unconditioned stimuli. When these acquire a certain strength, as a result of prolonged effortful behavior, an unconditioned resting response is elicited. This unconditioned resting response in a sense provides its own reinforcement since the decay of reactive inhibition occurs as a result of it. From the simplest principles of conditioning, it now follows that stimuli surrounding the S will come to evoke the resting response independently. This conditioned resting tendency is conditioned inhibition, ${}_sI_R$.

The identification of conditioned inhibition as a conditioned response means that, among others, it will have the following properties:

1. As a conditioned response, it will not be acquired in the absence of an unconditioned stimulus of sufficient strength to elicit the unconditioned response. What this means for motor learning is that I_R will have to attain a minimal value great enought to force the resting response for ${}_sI_R$ to develop. In support of this conjecture, the author was able to show (5) that no ${}_sI_R$ developed during the first 150 sec. of continuous practice on the inverted alphabet task.

2. Conditioned inhibition should increase as a negatively accelerated function of practice. The only evidence available on this point (5) confirms this postulate.

3. Conditioned inhibition should be relatively permanent; it should show little tendency to diminish with the mere passage of time.

Both conditioned inhibition and reactive inhibition have the effect of decreasing the level of the performance curve in learning. At this point it is probably worthwhile making explicit the point that I_R has now been assigned two somewhat different functions. As we have already seen, it is regarded as a drive which produces resting. Now we are saying that I_R depresses performance, by which we mean that it decreases work out-put, probably by interfering with efficient muscular coordination. This dichotomizing of the function of I_R is necessary to handle a postulated instance (and experimental results to show that it exists) in which performance is decreased by I_R alone without the development of ${}_sI_R$. Elaboration of this point is probably necessary. Suppose that we assume that the typical depression of the learning curve under conditions of massed practice occurs simply because the accumulation of reactive inhibition under these conditions produces rest-

ing frequently and these rests take time; that is, suppose we assume that performance is depressed only by the first of the functions which we have assigned I_R. Note that this formulation will handle the typical findings on studies of the distribution of practice. The greater the degree of massing of practice, the greater the amounts of reactive inhibition accumulated, the more time spent in resting behavior, and consequently the lower the level of performance. But this formulation also demands the development of conditioned inhibition during any practice series in which performance is less than optimal. Since I_R will dissipate during the rests, resting behavior is rewarded and conditioned inhibition will develop. Suppose now (to anticipate some of the results of the present experiment) that a long rest is introduced into one of these less than optimal learning situations and that, after the rest, performance equals that of a group practicing the same task under conditions of optimal distribution. Then we have evidence that, during the earlier practice, performance was inhibited by reactive inhibition alone and that no conditioned inhibition has developed. Theoretically, however, ${}_sI_R$ should have developed. It is to avoid this contradiction that we have assigned one of our major constructs two differ ent functions.

By contrast to our position that I_R needs two distinguishable properties, Hull (3) assigns it only the property of reducing the level of performance. Conditioned inhibition, he says develops when the stimuli present at the cessation of a response get conditioned to I_R. The main trouble with this postulate is that, within Hull's system, it is *responses* that are supposed to get conditioned to stimuli, not intervening variables like I_R. A more immediate dificulty with this position is that it seems to demand that ${}_sI_R$ being to develop whenever cessation of activity and reactive inhibition occur together. This contradicts the facts. Since I_R and ${}_sI_R$ depress performance curves, factors that minimize the effects of inhibition will facilitate performance. The most important of such factors is rest, during which reactive inhibition is known to dissipate. From the facilitation of performance resulting from the decay of I_R with rest, two important classes of experimental phenomena are deduced. These are (1) the advantages of distributed over massed practice in learning and (2) the phenomenon of reminiscence.

Purpose of the Experiment

The purpose of the experiment to be described in this paper was two-fold. First, it is considered as an empirical investigation concerned with the relationship between certain aspects of performance and the degree of distribution of practice. Second, it was designed as a test of some of the implications of the theory just described, especially of the hypothesis that the development of ${}_sI_R$ is contingent upon the prior development of a certain, threshold amount of I_R. Our main concern will be with three predictions based on the theory outlined in the preceding paragraphs. All of these predictions are about the results of an experiment in which a number of groups of *S*s practice a motor task under conditions involving trials of a constant length, but different degrees of distribution of practice. After a specified number of trials, each group is given a rest of several minutes and then a final practice trial. For this experiment, the theory permits the following deductions:

1. Prior to the rest, at any specified stage in learning, the level of performance will be a negatively accelerated function of the length of the time between trials. Since reactive inhibition dissi-

pates with rest, longer rest pauses which allow for greater decay of I_R will result in superior performance. At the same time, the curve which relates performance to the degree of distribution will be negatively accelerated because the decay function of reactive inhibition is exponential.

2. After the rest, there will be an increase in performance (reminiscence), the amount of which will vary inversely with the length of the time between trials. The amount of I_R dissipated between trials is greater in the case of the groups allowed the longer rest pauses during the pre-rest learning trials. Therefore, at any specified stage of learning, for these groups, there will be less I_R remaining to be dissipated and less reminiscence is possible.

The two limiting cases to this prediction are at least as important as the prediction itself. *First*, if any of the experimental groups are allowed rest pauses long enough to permit the complete dissipation of reactive inhibition, they will show no reminiscence effect. During learning, the performance curve will not have been depressed by any decremental factors and, therefore, no improvement after recovery from such factors is to be expected. *Second*, our postulation of a critical, threshold amount of I_R which sets a limit to the degree of development of this variable leads to the prediction that groups learning under high degrees of massing may show large and identical amounts of reminiscence, if two different degrees of massing produce this limiting value of reactive inhibition. Such groups would, however, differ in the amount of condition inhibition acquired.

3. After the rest, at least some of the groups will show a level of performance equal to that of the most highly spaced group. If it is necessary for I_R to reach some critical value for the development of ${}_SI_R$ to begin, and if the trials in the learning series in some conditions are separated by rests long enough to allow most of the reactive inhibition to dissipate, then it follows that no conditioned inhibition may ever develop, because the net amount of reactive inhibition will remain below the level required for the development of ${}_SI_R$ to begin. In this case, during learning performance is depressed only by the amount of the temporary I_R generated. Removal of this reactive inhibition by a rest means that all the decremental factors will have been dissipated. This in turn means that the post-rest performance of a group which has learned under conditions where I_R but no ${}_SI_R$ develops will be equal to that of a group learning under conditions where reactive inhibition itself was kept at a minimum by optimal distribution of practice trials.

One final question with which this paper is concerned cannot be put in the form of a specific prediction because the theory is not yet developed to the point where one can be made. This concerns the relationship between the *rate* of learning and the distribution of practice. So far as the author knows, there is in the literature only one serious attempt at the expression of this relationship. This attempt was made by Kientzle (4). She fitted equilateral hyperbolas to the performance data of her *S*s and used a rate measure based on these functions as her index of the speed of learning. When this rate measure is plotted as a function of the degree of spacing of practice trials, it is found to increase, but in an irregular manner, with the length of time between trials, so that the relationship between degree of spacing and rate of learning is hard to specify more definitely. One of the variables helping to confuse the issue is, we suspect, to be found in the rather complex relationship between performance and distribution early in practice. Kimble and Bilodeau (7) have shown that, during the early learning trials, the degree of distribution is not an important variable in determining performance scores on a rate of manipulation task. Kimble (6) subsequently showed that the same is true for the inverted alphabet. This means that, early in learning, rates of acquisition for several conditions of distribution will be indistinguishable. This in turn probably depends upon a fact sometimes overlooked in studies of distribution;

namely, that extreme degrees of spacing may, during the early practice trials, be less beneficial than a more moderate degree of distribution. Motor learning seems to involve the ability to adopt a perceptual or motor 'set' which is developed as a part of the acquisition of the motor skill. Until the skill of adopting this set rapidly is well learned, degrees of spacing greater than a few seconds are detrimental to performance because, during long rests, the motor set is lost and, when practice is resumed, time is required to 'warm up' to the learning task. As learning progresses, however, the ability to warm up rapidly is probably acquired to approximately the same degree by all *S*s regardless of the condition of distribution under which they are learning. When this happens, performance will depend more directly upon the spacing of practice since the effectiveness of this variable will be less obscured by the need to warm up. Because of these considerations, our analysis of the rate of learning will be one in which the early learning trials are excluded.

Experimental Procedure

The learning task used in the experiment was the alphabet printing task previously used by Kientzle (4) and by Kimble (5), in which the *S* prints the alphabet upside down, beginning at the right side of the page and working to the left. The experimental materials were identical with those previously described (5). The procedure was as nearly like that previously described as possible.

A total of five experimental groups were run. All were given 21 30-sec. trials under the following conditions:

1. The length of the rest pause between trials for the five groups was 0, 5, 10, 15, and 30 sec.

2. All groups except the 30-sec. group were given a 10-min. rest between the 20th and 21st trials. The 30-sec. group, which was a control group, received the last trial 30 sec. after trial 20.

The number of *S*s in each group is given in the last column of Table II.

Results

A. *Learning*

The mean score for each group for each of the 20 learning trials is presented in Table I. The data are presented in this way rather than graphically because of the great amount of overlapping in the data for approximately the first 10 trials. An examination of the mean scores will show that only the 0-sec. group differs from the others to any great extent during these trials. Part of the reason for the overlapping of the data will be apparent from an examination of Fig. 1. This figure was constructed to show the level of performance as a function of the degree of spacing of practice at four different stages in learning. Plotted in these four curves are the mean scores for each of the five groups on trials 5, 10, 15 and 20. Our immediate concern is with the lower curve plotted to show the performance of each group on trial 5 as a function of the duration of the inter-trial rest. Here we note that,

TABLE I

Mean Score on Each Trial for Each Experimental Group

Group	Trial																			
	1	2	3	4	5	6	7	8	9	10	*11	12	13	14	15	16	17	18	19	20
0 sec.	18.4	14.3	17.1	17.5	20.0	19.5	21.0	20.2	21.1	22.0	20.1	21.5	23.2	22.1	23.0	23.4	23.6	23.1	23.9	24.9
5 sec.	19.9	20.3	22.3	23.2	24.2	24.9	24.7	26.2	26.9	26.6	24.8	26.2	27.3	27.7	28.0	28.6	28.3	28.6	29.6	29.5
10 sec.	19.8	20.5	21.3	22.8	24.1	25.2	25.5	25.9	25.7	26.2	24.8	26.9	27.3	28.6	28.7	28.7	29.4	29.4	30.3	30.1
15 sec.	19.3	20.8	21.1	22.7	23.9	24.9	25.9	26.0	26.0	27.2	26.6	27.5	29.2	28.9	30.0	30.1	30.3	31.1	31.3	31.9
30 sec.	17.7	19.3	21.3	23.2	22.8	24.3	26.0	27.4	27.8	28.6	28.7	30.1	30.8	31.4	31.2	32.5	33.3	33.7	34.0	34.6

* The decrease in performance occurring on trial 11 results from the fact that the experimental procedure required the turning of a page after trial 10.

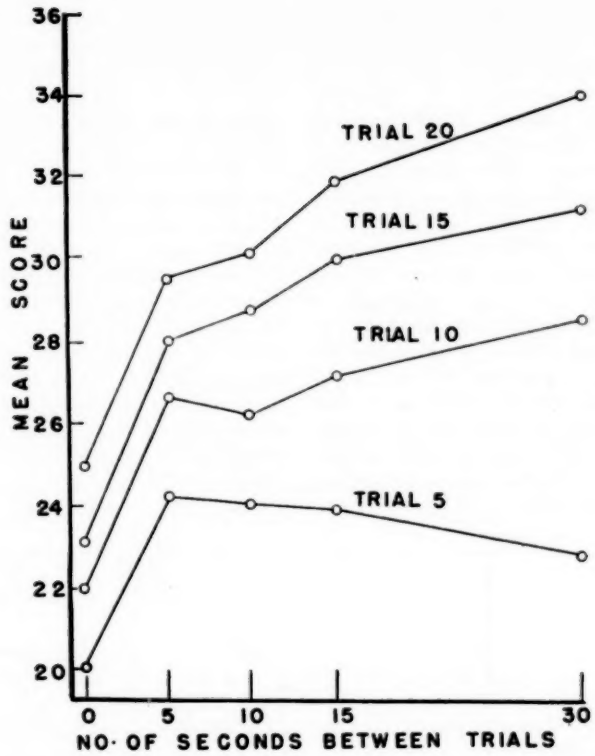

FIG. 1. Performance as a function of the number of sec. between trials on trials 5, 10, 15 and 20. Points plotted are mean scores of each group on the trial indicated on the individual curves.

at this stage of practice, five sec. is the optimal inter-trial interval, and that either shorter or longer intervals produce inferior performance. One way in which this result can be interpreted is on the assumption that a part of the skill involved in the alphabet writing task is a perceptual-motor adjustment which is lost during the longer rest-pauses between trials. When this happens the advantages of spaced practice are obscured. One would expect this detrimental effect of spaced practice to occur early in learning before the *habit* of adopting the correct adjustment is well learned. From trial 10 on, the curves in Fig. 1 are more nearly like what was predicted. In general, they are negatively accelerated and seem to be approaching an asymptotic level at a point somewhat beyond 30 sec. on the abscissa. In other words, these curves show that performance is a negatively accelerated function of the degree of distribution of practice.

Let us now turn to an analysis of the rate of learning, or of the speed with which the motor habit is acquired. An examination of Table I will serve to show that, for about the first 10 trials, the scores of the various groups are quite similar, and that, for these trials, the speed of learning

does not vary. As has been suggested, this fact is probably the result of two opposed processes attached to spaced practice early in learning. By trial 12, however, the five experimental groups have become fairly well separated. Data from this trial on to the end of the learning session are plotted in Fig. 2. The lines through the empirical points are best-fitting straight lines obtained by the method of averages. Although a straight line would not fit the data for the whole 20 trials, such a function does seem to describe the results for this segment of the trials quite adequately. One notable feature of these curves is the progressively steeper slope of the straight lines as a function of the length of the inter-trial interval. This fact is presented more clearly in Fig. 3, where the slope constants of the curves in Fig. 2 are taken as indices of the speed of learning and plotted as a function of the duration of the inter-trial rest. The result is an increasing function which is well-fitted by another method of averages straight line.

The rate of learning has now been shown to bear a complex relationship to the length of the inter-trial rest. Early in learning, the length of the rest pause does not affect the rate of learning in any consistent way, perhaps because lengthening the rest at this stage of practice has two opposed effects which tend to cancel one an-

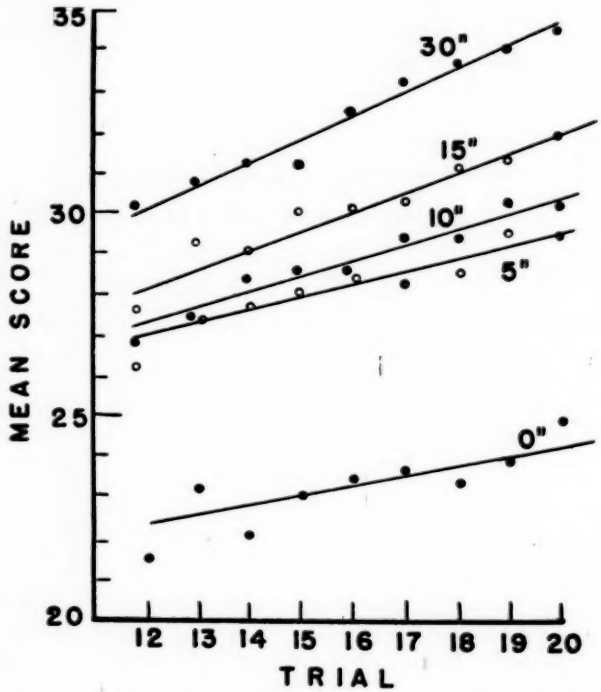

FIG. 2. Learning curves from trial 12 to the end of the learning session. The lines through the empirical points are best fitting straight lines obtained by the method of averages.

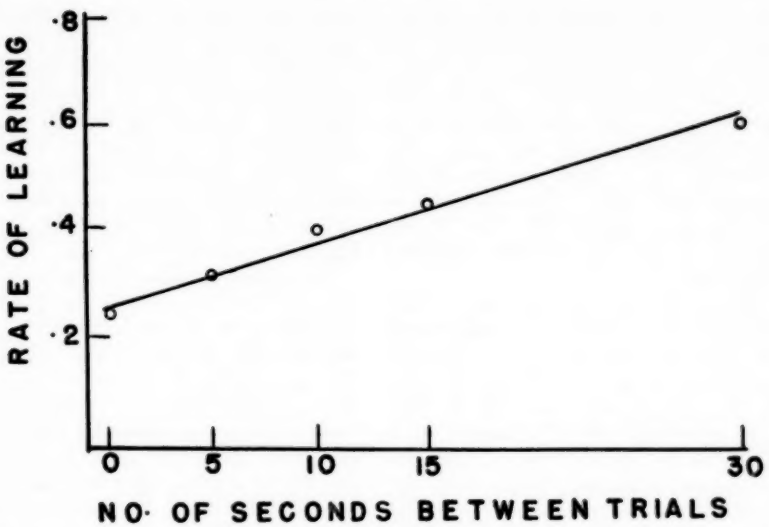

FIG. 3. Final rate of learning as a function of the degree of spacing of practice. The rate measure plotted is the slope constant of the individual curves plotted in Fig. 2. The straight line is a line of best fit by the method of averages.

other. Later on in learning, the speed of acquisition of the motor skill is an increasing linear function of the degree of distribution of practice.

B. *Reminiscence*

All of the important data of reminiscence are presented in Table II. In successive columns of this table, the data presented are:

1. The absolute amount of increase after the 10-min. rest.
2. A critical ratio, which tests the hypothesis that the population increase is zero.
3. The score for each group on trial 21, after the 10-min. rest.
4. The amount by which this score falls below the trial 21 score for the 30-sec. group.
5. A critical ratio testing the hypothesis that this difference (column 4) is zero.
6. The number of *S*s in each group (*N*).

TABLE II

DATA ON REMINISCENCE AND POST-REST PERFORMANCE

Group	Gain	*CR*	Score on Trial 21	Diff. between Trial 21 Score for 30-Sec. Group and Every Other Group	*CR*	*N*
0 sec.	3.8	8.2	28.8	4.6	4.0	65
5 sec.	3.9	6.6	33.4	0.0	0.0	39
10 sec.	2.5	4.3	32.6	.8	.6	39
15 sec.	.8	1.6	32.7	.7	.6	35
30 sec.	—	—	33.4	—	—	46

It will be noted that the two groups with the shortest inter-trial rest pauses showed essentially the same amount of reminiscence, and that the groups with longer rests between trials showed progressively less gain, presumably as a result of the longer interpolated rest. The amount of reminiscence shown by the 15-sec. group is not significantly different from zero.

With respect to the comparison between the reminiscence point (score on trial 21) and the 30-sec. group's performance on this trial, it is only in the case of the 0-sec. group that this difference is significant. This leads us to the conclusion that no permanent decrement in performance results from the massing of learning trials unless this massing is complete.

Summary and Discussion

The results of the present experiment that seem important for a comprehensive theory of motor learning are the following:

1. No difference in the amount of temporary reactive inhibition as indicated by the amount of reminiscence was obtained for the groups that learned under the two conditions with the greatest degrees of massing of practice, in spite of a great difference in the level of performance at the end of the learning session and prior to the rest.

2. Except for the case noted above, the amount of reminiscence and, therefore, the amount of reactive inhibition, decreases as the inter-trial interval increases.

3. There is no permanent decrement in performance unless the inter-trial rest is reduced to zero or nearly that.

4. Except in the case of the 0-sec. group, the effects of differences in the length of the rest pause do not appear early in learning, at least for this particular task.

5. After the early trials, the level of performance is a negatively accelerated function of the length of the inter-trial rest pause.

6. When differences in the *rate* of learning finally appear, they occur as an increasing linear function of the length of the inter-trial rest.

In the main, these findings support the predictions made on the basis of our theory as it was outlined in the introduction to this paper. Among the more critical results is the fact that no difference in the amount of reminiscence was shown by the 0-sec. and the 5-sec. groups. Since the amount of reminiscence is thought to be correlated with the amount of reactive inhibition, this is considered as confirmation of the portion of the theory which maintains that there is a crictical amount of reactive inhibition that an individual will tolerate. The fact that these groups showed the same amount of reminiscence would indicate that the same amount of I_R had accumulated in each case. Yet the 5-sec. group learned under conditions in which the dissipation of I_R between trials was possible. No such opportunity for the dissipation of inhibition was afforded the 0-sec. group. The fact that the same net amount of I_R seems to have developed under these two different conditions suggest that there is some limiting value of I_R beyond which this variable does not develop. In the case of the 0-sec. group, it is evident that *conditioned* inhibition developed and was playing an important part in the depression of the performance curve during learning. Evidence for the presence of ${}_sI_R$ is to be found in the fact that, after the rest, the performance of this group is still considerably below that of the 30-sec. group. This permanent

decrement is considered as evidence for conditioned inhibition.

Of particular importance seems to be the fact that as little as five sec. between trials prevents the development of ${}_{s}I_{R}$. Our prediction had been merely that *some* of the groups would develop no conditioned inhibition in spite of the fact that their performance was depressed by the accumulation of reactive inhibition. From the experimental results it appears that only the most highly massed practice conditions favor the development of ${}_{s}I_{R}$. Since such practice conditions are not commonly encountered in learning experiments or elsewhere, it would seem that conditioned inhibition plays a much less important part in behavior than had previously been thought. In this connection, a particularly important point arises when one considers the postulated role of conditioned inhibition in the accomplishment of experimental extinction in the conditioning studies. The permanent aspects of experimental extinction, according to Hull, are assignable to the development of ${}_{s}I_{R}$ during the extinction series. Since conditioned inhibition is a habit, there will be no spontaneous recovery of an extinguished conditioned response to the extent that the response has been inhibited by ${}_{s}I_{R}$. However, the degrees of distribution of practice used in the extinction of conditioned reflexes are such that probably no ${}_{s}I_{R}$ could be developed if the requirements for the development of conditioned inhibition are the same in the conditioning situation as they are in motor learning. This means either (1) that the conditioned inhibition postulated by Hull to develop during the extinction of conditioned responses is fundamentally different from that developed in motor learning, or (2) that the concept of experimental extinction as depending entirely upon inhibitory factors is wrong and that the habit itself is being diminished during extinction.

The finding that the amount of reminiscence generally decreases as a function of increases in the length of the inter-trial rest was predicted on the basis of the assumption that during practice greater amounts of I_R will be dissipated during the long rest pauses, with the result that performance will be less affected and that there will be less reminiscence possible. Somewhat surprising, however, is the finding that, with as little as 15 sec. between trials, the amount of reminiscence is reduced nearly to zero. This means that, with practice trials 30 sec. long, 15 sec. between trials is near the upper limit of favorable degrees of distribution. A number of important questions for the understanding of cyclical motor learning require a comparison of these results with those obtained when the practice trial is longer.

Closely related to the decrease in the amount of reminiscence with greater degrees of spacing is the negatively accelerated, increasing form of the functions relating the level of performance to the length of the rest-pause which were presented in Fig. 1. These curves were about as they had been expected to be on the basis of theoretical considerations, except in the case of the one constructed for the five experimental groups on the fifth trial. Here the descending level of performance beyond the point plotted for the 5-sec. group is contrary to the prediction made on the basis of the characteristics of the inhibitory process alone. What seems to be necessary in order to understand this phenomenon is the concept of some adjustive habit which *S* learns to make more and more rapidly as learning

progresses. Some postural or perceptual adjustment is perhaps involved. The importance of some such factor as this is, of course, already recognized by investigators in the field of motor learning. Ammons (1) makes room for such events in his theory of motor learning, and has done the best job to date of relating the warm-up phenomenon to the independent variables in motor learning (2).

Our findings on the *rate* of learning as a function of the degree of spacing represent only a first approximation. What we have done is merely to consider the rate of learning as a function of the degree of spacing in the portion of the acquisition curve where the function seems to be least complicated. It seems proper to exclude the early trials in such an analysis because of the complications introduced into the problem by the initially detrimental effects of the wider degrees of distribution. When this was done, the relationship between degree of spacing and rate of learning seemed to be well described by a straight line. Whether or not the same relationship would be found had more learning trials been available on which to estimate the rate of learning or had wider degrees of distribution been used is an empirical question that can be answered only on the basis of a more extensive experiment.

(Manuscript received May 24, 1948)

References

1. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281.
2. Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance before and after a single rest. *J. exp. Psychol.* 1947, **37**, 393–411.
3. Hull, C. L. *Principles of behavior.* New York: D. Appleton Century, 1943.
4. Kientzle, M. J. Properties of learning curves under varied distributions of practice. *J. exp. Psychol.*, 1946, **36**, 187–211.
5. Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.
6. Kimble, G. A. A further analysis of the variables in cyclical motor learning. *J. exp. Psychol.*, 1949, **39**, 332–337.
7. Kimble, G. A., & Bilodeau, E. A. Work and rest as variables in cyclical motor learning. *J. exp. Psychol.*, 1949, **39**, 150–157.
8. Kimble, G. A., & Horenstein, B. R. Reminiscence in motor learning as a function of length of interpolated rest. *J. exp. Psychol.*, 1948, **38**, 239–244.

EXTINCTION AS A FUNCTION OF PARTIAL REINFORCEMENT AND DISTRIBUTION OF PRACTICE [1]

BY VIRGINIA F. SHEFFIELD

Yale University

PURPOSE

The purpose of this study was to determine whether distribution of practice influences the effect of partial reinforcement on resistance to extinction. The problem has theoretical significance because of its bearing on a general hypothesis concerning extinction which may explain the finding that partial reinforcement results in greater resistance to extinction than reinforcement on every trial.

For some time it has been recognized that reinforcement on every trial is not essential either for the establishment or for the maintenance of a conditioned response. Pavlov (18) describes an experiment performed in his laboratory in which conditioned salivation in a dog was established with food reinforcement on every other trial and on every third trial. Others (2, 3), while not specifically studying partial reinforcement, have also obtained successful acquisition using partial reinforcement instead of reinforcement on every trial. Also it has been shown (1) that a conditioned response, established with reinforcement on every trial, may be maintained at almost its original level with relatively infrequent reinforcement.

The most striking finding with partial reinforcement, however, is that not only are conditioned responses readily acquired with partial reinforcement, but also they are considerably more resistant to extinction than when reinforcement is given on every trial. For example, Skinner (20), using bar pressing with rats, interspersed reinforcements at regular time intervals ('periodic reconditioning') or after given numbers of responses ('reinforcement at a fixed ratio'). He found that many more responses were required to produce extinction than when reinforcement had been given for each response.

Attention was particularly focused on the phenomenon of increased resistance to extinction following partial reinforcement in an experiment by Humphreys (11) using the conditioned eyelid response in human subjects. Of two groups of subjects, one trained with reinforcement on every trial and one trained with reinforcement 'randomly' on only half the trials, the latter group responded at a significantly higher level throughout extinction. Humphreys (13) later obtained the same result using the conditioned galvanic response in humans.

It has been claimed (8, 11) that the increased resistance to extinction with partial reinforcement challenges the adequacy of stimulus-response learning theory. Humphreys (11) proposed the principle of *expectancy* as an alternative theory. This 'principle' is not rigorously stated and depends on commonsense concepts in its explanation of the results. The 'expectancy' explanation may be paraphrased as follows. Conditioned responses are the consequence of the subject's expectation that reinforcement will appear. In extinction after reinforcement on every trial, the response rapidly disappears because the sudden shift from uniform reinforcement to uniform nonreinforcement makes it easy to

[1] This article reports part of a dissertation submitted to the faculty of the Department of Psychology of Yale University in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Dr. Neal E. Miller, under whose direction the research was conducted, and to Dr. Clark L. Hull and Dr. Carl I. Hovland, who served on the advisory committee.

change to an expectation of uniform nonreinforcement. But in extinction after partial reinforcement, the subject continues to expect that reinforcement will be periodic as it was during training—extinction is prolonged by his expectation that reinforcement will be reintroduced. This interpretation is contrasted with stimulus-response learning theories, which usually assume, in one form or another, that reinforcement strengthens a conditioned response and nonreinforcement extinguishes it. This suggests that the infrequent reinforcements and the interspersed nonreinforcements which characterize partial reinforcement would produce a weaker response, which would extinguish more rapidly than would be the case with uniform reinforcement.

However, it is believed that ordinary stimulus-response learning concepts, such as those used by Guthrie (7) and Hull (10), can explain the increased resistance to extinction. The present study was designed to test a theoretical interpretation which utilizes only stimulus-response concepts. The interpretation hinges on a hypothesis about a general factor in extinction which is particularly applicable in the case of partial reinforcement.

Theory

The basic hypothesis is that extinction necessarily involves different cues from those used during training. Omission of reinforcement alters the context and makes extinction a case of 'transfer of training' in which a certain amount of generalization decrement is expected because of the change in cues. This hypothesis would be applied in explaining the effect of partial reinforcement on extinction as follows. Occurrence of reinforcement on a given trial produces effects which, in varying degree, provide characteristic stimuli at the start of the following trial, these stimuli becoming part of the total stimulus pattern acting at the start of the next trial. When reinforcement is given on every trial, the after-effects of the reinforcement will be part of the conditioned stimulus pattern on every trial after the first. In the present experiment, where rats were trained to run in an alley to food, the after-effects of reinforcement would perhaps include continued food taste and even particles of food in the mouth, temporary relaxation, etc. When extinction is begun, the stimulus pattern is changed not only by the absence of the after-effects of reinforcement, but also by the presence of whatever new stimulation results from the absence of reinforcement, particularly from the subject's reaction to that absence. In the present experiment, the after-effects of nonreinforcement would include such new reactions as searching, conflict, frustration, slowing up, etc. This change in the conditioned stimulus pattern should result in a weakening of the conditioned response, because the cues present during extinction are at a point on the generalization gradient at which responding is expected to be weaker than to the cues present during training.

However, when training with partial reinforcement is given, the subject is exposed, on reinforced training trials that follow nonreinforced trials, to cues which are normally present only during extinction. Examples of such cues for animals reinforced with food would be lack of food particles in the mouth, a state of frustration, etc. In other situations, examples would be the absence of irritation from recent air puff to the eye or shock to the wrist, cues from any relaxation resulting from freedom from the air puff or shock, implicit verbalizations relating to nonreinforcement, etc. With these nonreinforcement cues as part of the current stimulus pattern, reinforcement is

reintroduced, and the subject therefore learns to perform the response in the presence of such nonreinforcement cues. Since the response has become conditioned during training to the cues characteristic of extinction, there is less loss through a change in the conditioned stimulus pattern when reinforcement is withdrawn completely than is found after training by reinforcement on every trial.

Thus, the initiation of extinction trials produces a relatively large change in the conditioned stimulus pattern when it follows training with reinforcement on every trial, but much less change when it follows training with partial reinforcement. Because the response evoked by a generalized stimulus is weaker than that evoked by the reinforced stimulus, the conditioned response during extinction will be weaker in the former case than in the latter. The possibility of the operation of this mechanism has been recognized by other authors (4, 9, 16, 17).

Method

The proposed test of this interpretation consisted of controlling some of the after-effects of reinforcement and nonreinforcement by spacing of trials. The assumption was that if animal subjects were used with trials widely spaced, most of the after-effects of reinforcement or nonreinforcement would have dissipated by the start of the next trial, making the conditioned stimulus pattern much the same whether reinforcement had or had not been received on the preceding trial. The group receiving reinforcement on every trial and the group receiving partial reinforcement would thus be trained with more similar conditioned stimulus patterns than when trials were massed. Similarly the change in conditions when extinction was started would be more nearly the same for the two groups; as a result, the difference in resistance to extinction should be reduced.

Specifically, rats were trained to go down an alley with food at the end of the alley as reinforcement. Such after-effects of reinforcement as food taste and such after-effects of nonreinforcement as frustration may reasonably be assumed to dissipate with the passage of time and the performance of other behavior. Therefore, it follows that the extent to which differentiable reinforcement and nonreinforcement cues enter into the total stimulus pattern on a given trial will be maximized if trials are spaced close together and will be minimized if trials are spaced far apart. Massing of training trials, then, should give the maximum advantage to partial reinforcement in preventing extinction as compared with reinforcement on every trial, whereas relative spacing of trials should diminish or destroy this advantage. It is this deduction which the present experiment was designed to test.

Seventy-two rats were used in the experiment. Half the animals were trained with 100 percent reinforcement for 30 trials, half with reinforcement randomly on 50 percent of the 30 training trials. Of each group, half were trained with the trials massed (15-sec. interval), half with the trials spaced (15-min. interval). The question to be answered was: would partial reinforcement have less advantage in resistance to extinction in the latter group?

It seemed probable that the spacing of trials, both during training and during extinction, would have the direct effect of increasing resistance to extinction as well as its indirect effects through interaction with partial reinforcement. At the same time, it seemed probable that a change from massed to spaced or from spaced to massed when extinction begins would have the effect of hastening extinction. In order to balance completely for these factors, each of the four training groups was divided for extinction, half receiving massed extinction trials and half spaced extinction trials. This procedure not only served as a balancing procedure in the main problem to be investigated, but also had intrinsic value in supplying data on the problem of the relative resistance to extinction when extinction trials are massed compared with when they are spaced. The results of this aspect of the experiment will be presented in a separate article (19). Table I shows the design in schematic form; each cell of the table represents a separate subgroup of the experiment.

Details of Apparatus and Procedure

The apparatus, shown in Fig. 1, consisted of a four-foot alley connecting a starting box and a goal box. Walls throughout were nine inches high. The interior of the starting box was painted white, and the interiors of the alley and goal box were painted black. The light in the room was from a ceiling fixture directly over the center of the apparatus. The two doors were

TABLE I

DESIGN OF THE EXPERIMENT

	Massed Training		Spaced Training	
	Massed Ext.	Spaced Ext.	Massed Ext.	Spaced Ext.
100% Reinf.	100% MM 9	100% MS 9	100% SM 9	100% SS 9
50% Reinf.	50% MM 9	50% MS 9	50% SM 9	50% SS 9

operated vertically by strings and pulleys. Small pellets of wet mash, made from powdered Purina dog chow, flour, and water, were presented in a white food cup in one corner of the goal box. On all nonreinforced trials, both during partial reinforcement and during extinction, the white food cup was removed from the goal box. A black curtain placed diagonally at the entrance to the goal box cut off sight of the food cup from the alley, so that an animal could not tell whether the cup was present or absent until he was actually inside the goal box.

The rats were always placed in the starting box facing the closed door. After approximately two sec. the door of the starting box was raised. When the rat left the starting box, the door was lowered; when he stepped onto the floor of the goal box, the door near the goal box was lowered to prevent retracing. Microswitches mechanically operated by the animal's weight on hinged floor sections in the two end boxes activated Springfield timers which recorded starting time (time from opportunity to respond until departure from the starting box) and running time (time from leaving the starting box to entering the goal box). The starting box was covered with a glass lid to prevent animals from prematurely operating the timer switches by jumping into the air before they actually left the box. The alley and goal box had no covering. A holding relay prevented re-operation of the switches by retracing. The goal box door was set back a short distance in the alley so that once the animal's forepaws were inside the goal box and the second timer had stopped, the door could be closed even if the animal did not immediately enter the box completely.

Seventy-two male albino rats from the albino farms at Redbank, N. J. were used, their ages ranging from 83 to 107 days (average 93 days) at the time they were started in the experiment. They had never been used in any previous experimentation.

Before and during the experiment the animals lived in the experimental room. During the two weeks prior to training, the animals were given daily taming sessions, consisting of handling during feeding periods until all would eat pellets of food from the experimenter's hand. They were also given practice at eating in unfamiliar surroundings—cages unlike their living cage, a table top, and a large cardboard carton. During these two weeks, they were fed a restricted diet of wet mash given about noon. Two days prior to the start of the experiment, the feeding time was shifted to the time at which their training would be started, a shift ranging from zero to four hours.

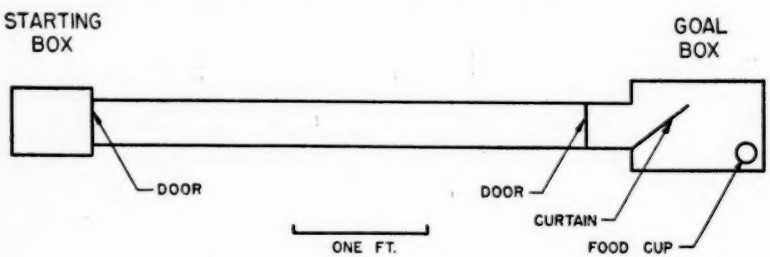

FIG. 1. Apparatus

On the first day of the experiment proper, all animals were given 10 pre-training trials in the apparatus under constant conditions of massed trials and 100 percent reinforcement. This pre-training provided a criterion for eliminating, in advance of the application of the experimental variables, those rats too timid to explore the alley (and probably also helped insure that the rats scheduled to receive partial reinforcement would not give up on the first few training trials). If during the course of the 10 pre-training trials an animal took more than five min. to find the food or to eat it, he was discarded. On this basis, 14 rats were eliminated. Two rats were discarded after pre-training despite acceptable performance in the alley, one because he violently resisted being picked up and because he consistently bit the experimenter, the other because the apparatus broke down during experimental training. All discards were replaced to make the final group of 72 experimental animals.

The 10 pre-training trials were given on the first day; 30 training trials were given on the second day; and 30 extinction trials were given on the third day. The trials were started at the same time on each of the three successive days. The 50-percent reinforcement groups, having the same number of training trials as the 100-percent reinforcement groups, received only half as many reinforcements during training. At the end of their training trials, when the animals were given the balance of their ration for the day, this difference in amount eaten between 50-percent and 100-percent reinforcement animals was made up.

At the end of a run, animals were left in the goal box for 10 sec., or until they found and ate the food, longer times being necessary on early training trials. Ten sec. was an adequate time to eat the pellet of wet mash once the animal had learned, and this time was used whether food was present or not, during pre-training, training, and extinction. The animals were then removed to their individual living cages to await the next trial. Animals receiving massed trials were left in their living cages for 15 sec. between trials. This amount of time was required to record the time measures and reset the apparatus for the next trial. Due to variable running time, the time from the start of one trial to the start of the next was thus variable, but it very rarely ran over two min. and was usually close to ½ min. except on the first few training trials and during extinction. For spaced trials, the time from the start of one trial to the start of the next was 15 min. During extinction the animals were removed from the apparatus if either the starting time or the running time ran to two min., but all animals were given the full 30 extinction trials.

Since the training session extended over a much longer time for spaced groups than for massed groups, it was not possible, with the experimental setup used, to obtain strictly comparable hunger drive from group to group. However, the hunger drive was kept at a high enough level for all rats that it was felt that any variability in absolute level would be unimportant. In any case, whatever drive difference may have been present between certain of the groups did not enter into the basic comparisons of 50 percent and 100 percent reinforcement, for which drive was constant in groups compared. A restricted diet of seven gm. (dry weight) of the dog chow and flour mash was used before and during the experiment. Thirty pellets, the most any rat received during training, weighed about 2¼ gm. (dry weight), which was only one-third of the restricted diet and an even smaller fraction of what any rat would eat if allowed an unlimited supply of food. Rats given massed training trials had a 23-hour hunger drive at the start of extinction. Spaced training was spread through a period of about 7¼ hours, at the end of which time the animals were given the balance of their food for that day. Extinction for space-trained animals was started 24 hours after the start of training, i.e., 16¾ hours after feeding. In an attempt to make the hunger drive of space-trained animals more comparable at the start of extinction with that of mass-trained animals, a total food ration of only six gm. was given at the end of training instead of the usual seven gm.

In preparing the orders of reinforced and nonreinforced trials for presenting 15 rewards distributed over 30 training trials, certain restrictions were used. These restrictions were selected to produce orders which contained single nonreinforced trials, and successions of two in a row, three in a row, and four in a row, which closely approximated, within *each* series of 30 trials, the distribution expected by chance in an infinite series. In other words, chance frequencies were stratified within orders. Nine orders were prepared, each order being used once in each subgroup. In each of the orders prepared:

1. Four successive nonreinforcements and three successive nonreinforcements each appeared once.

2. Three successive reinforcements appeared twice.

3. Two successive reinforcements and two successive nonreinforcements each appeared twice.

4. Isolated single reinforcements appeared

five times, and isolated single nonreinforcements appeared four times.

In order to prevent partial-reinforcement animals from refusing to run in the early part of training and in order to end their training with the response at high strength, the following additional restrictions were used:

5. The first training trial was always rewarded.

6. The set of four successive nonreinforcements never appeared in the first half of training.

7. The last training trial was always an isolated single reinforcement.

Within these restrictions the distributions used were prepared from Tippett's (21) random numbers. The nine orders used were assigned randomly to the nine animals in each subgroup.

Statistical Analysis

The statistical analysis of the results treated the data as nine replications of the experiment, with eight degrees of freedom for each comparison made. This treatment was partly dictated by a balancing procedure used in assigning rats to the eight conditions.

The rats used were obtained in three different shipments, each shipment upon arrival being broken down into sets of eight cage-mates. Differences in ease of taming were apparent between the shipments and between the sets of eight cage-mates within each shipment (perhaps due to the more docile being removed from the shipping crate first). These differences were balanced out by dividing the rats from each shipment equally among the eight experimental conditions and by assigning one each of the eight rats in a cage to the eight experimental conditions. The assignment of the eight rats to the eight experimental groups was done randomly, using Tippett's tables of random sampling numbers (21).

As a result of this balancing procedure, the eight rats within a replication were more homogeneous than the rats within a given experimental group. The greater precision in the experiment obtained by this procedure was taken into account in the statistical analysis by basing comparisons on the eight degrees of freedom provided by the nine replications. Thus in testing for the significances of the mean differences between two major groups, the appropriate difference was found among rats within each replication, and the mean difference was tested from the distribution of nine differences so obtained. This procedure removed from the estimated variance of mean differences any correlation due to balancing of cage differences.

Results

Although two time measures were recorded—starting time and running time—comparison of the learning and extinction curves for the eight experimental groups revealed that the pattern of results was the same for the two measures. Therefore, the two measures were added together to make a single measure of 'response time' in order to increase the stability of the measurements and simplify the statistical analysis.

Acquisition

In Fig. 2 are shown learning curves for the four training groups. The

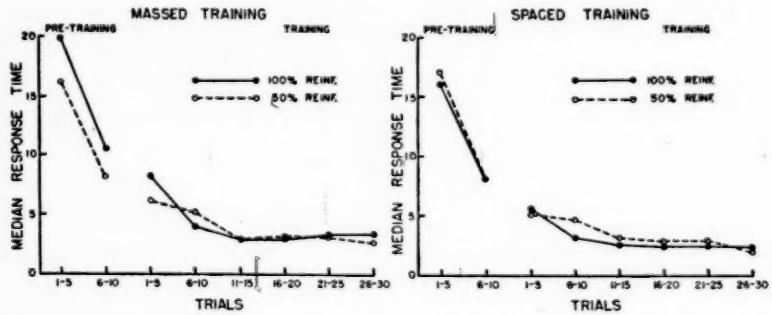

FIG. 2. Acquisition curves for the four training groups, showing median response time (averaged for groups of five trials) during the 10 pre-training trials and the 30 training trials

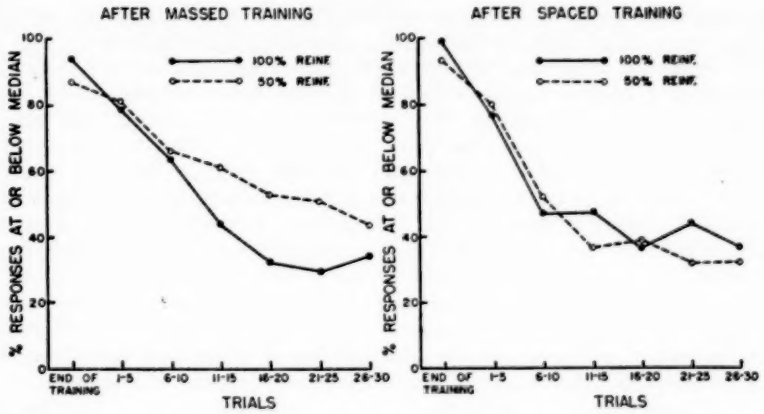

FIG. 3. Extinction curves for 100-percent and 50-percent reinforcement groups after massed training and after spaced training, showing percentage of responses in sets of five trials which were at or below 16.3 sec. (the overall median response time). The reference points at the beginning are the values for the last five training trials.

massed and spaced extinction groups have been combined for each training condition, since up to the end of training there was no difference in their treatment. The measure used is median response time, which was obtained for each trial and averaged for groups of five trials. Medians were used as more representative of the data because means would be unduly affected by occasional trials on which an animal had an unusually long response time. Pre-training and training are both shown.

Tests of the differences in the level of performance reached on the last half of the training trials show that there were no significant differences among the four groups. Groups receiving 100 percent reinforcement were very slightly superior to those receiving 50 percent reinforcement, but the probability that a difference this large and in this direction could arise by chance is .20 for massed training, .12 for spaced training, and .15 overall.[2] Groups receiving spaced training were very slightly superior to those receiving massed training, but the probabilities are .13 for 100 percent reinforcement, .28 for 50 percent reinforcement, and .15 overall.

Extinction

Fig. 3 shows extinction curves for the 100-percent and 50-percent reinforcement groups that were mass trained compared with the 100-percent and 50-percent reinforcement groups that were space trained (massed and spaced extinction being combined for each training group). Median response time could not be used in

[2] The measure used in computing significances of differences was the time, for each *animal*, on his median *trial* out of the last 15 trials. The measure used in the curves was the time, for each *trial*, of the median *animal* (groups of five trials being averaged). This difference between the methods of obtaining the two measures probably accounts for the fact that the slight superiority of 100 percent reinforcement during massed training indicated by the significance test is not apparent in the curves.

plotting extinction data since on some trials more than half the animals were removed from the apparatus because of exceeding the two-minute time limit, making the median indeterminate. Median response time over an individual animal's trials also could not be used as a measure of individual animals' performance in the analysis of extinction data because some animals exceeded the time limit on more than half of their extinction trials. A frequency measure was indicated, and that chosen was the number (or percentage) of trials with response times at or below the combined median response time for whatever groups were being compared. The median was chosen as the cut-off point partly because of its arbitrary nature and partly because it places the mean of the scores close to the middle of the range, which should give maximum sensitivity in comparing groups and at the same time avoid skewed distributions. The median time for all 72 animals on all 30 extinction trials was 16.3 sec., and it is the mean frequency of responses at or below this value for each set of five trials that is plotted in Fig. 3. The first point on the curves is the corresponding percentage value for the last five training trials, which is shown as a reference point.

Extinction after massed training.—As can be seen in Fig. 3, the results after massed training confirm the findings of previous investigators that training with 50 percent reinforcement results in greater resistance to extinction than training with 100 percent reinforcement. The results of a significance test are shown in Table II. All groups reached approximately the same level of performance at the end of training, and performance was also similar on the early extinction trials, but the two groups diverged after the first ten extinction trials. The significance test was made separately for all extinction trials and for the last half of the trials. The means in Table II for all 30 trials used the median of 14.2 sec., obtained for the 30 extinction trials of the 50-percent and the 100-percent mass-trained groups combined, as the cut-off point in getting each animal's score. The means for the last 15 trials used the median of these 15 trials, 24.4 sec., as the cut-off point. It should be pointed out that with the scoring procedure used, the potential range of scores is the same as the number of trials involved,

TABLE II

MEAN NUMBER OF RESPONSES MEETING CRITERION* DURING EXTINCTION AFTER MASSED TRAINING

	Mean Score		Mean Diff.	σ_M diff.	t	df	P**
	50% Reinf.	100% Reinf.					
All 30 ext. trials (Md. = 14.2 sec.)	17.1	13.0	4.1	1.87	2.2	8	.03
Last 15 ext. trials (Md. = 24.4 sec.)	9.0	6.0	3.0	0.72	4.2	8	<.01

* I.e., mean number of responses at or below median response time for the particular groups and number of trials being considered.

** For one tail of distribution.

TABLE III

MEAN NUMBER OF RESPONSES MEETING CRITERION * DURING EXTINCTION AFTER SPACED TRAINING

	Mean Score		Mean Diff.	σ_M diff.	t	df	P**
	50% Reinf.	100% Reinf.					
All 30 ext. trials (Md. = 19.1 sec.)	14.7	15.4	−0.7	2.71	0.3	8	.77
Last 15 ext. trials (Md. = 27.0 sec.)	6.9	8.0	−1.1	1.49	0.7	8	.50

* I.e., mean number of responses at or below median response time for the particular groups and number of trials being considered.

** For both tails of distribution.

so the means shown for all extinction trials and for the last half of the trials cannot be directly compared.

Table II shows that animals trained with partial reinforcement made a significantly larger number of responses during extinction which were below the median response time than did animals trained with 100 percent reinforcement, their superiority being more marked during the latter half of extinction than over all extinction trials.

Extinction after spaced training.—When training trials were spaced 15 min. apart, greater resistance to extinction following 50 percent reinforcement as compared with 100 percent reinforcement was absent. This result can be seen in Fig. 3. A significance test is shown in Table III. As in Table II, results are shown for all extinction trials and for the last half of the extinction trials. The scores on which Table III is based used the medians of the 50-percent and 100-percent space-trained groups combined for the relevant number of trials. The theory says nothing about the direction of this difference; as trials are more and more widely spaced, the difference should become smaller and smaller, but at what interval it would become zero and whether it would become negative are questions that would have to be answered empirically. For this reason, the *P*'s shown are for a difference this large in either direction.

Table III shows that with training trials spaced, animals trained with partial reinforcement did not differ significantly during extinction from animals trained with 100 percent reinforcement. The slight difference obtained was in the reverse direction from that obtained with massed training, 100 percent reinforcement being superior.

Effect of distribution of practice on extinction.—The results confirm the expectation from the stimulus-response analysis of partial reinforcement that the difference in resistance to extinction between groups trained with 100 percent reinforcement and groups trained with 50 percent reinforcement would be greater if the training trials were massed than if they were spaced. Tables II and III show that with massed training partial reinforcement significantly retarded extinction, whereas with spaced training it did not. They do not show, however, whether this obtained difference in the effect of partial reinforcement was significant. A significance test on this point is shown in Table

IV. Here the overall median response time of 16.3 sec. for all groups combined was used to obtain the scores over the 30 trials, and the median performance of 25.3 sec. for all groups combined on the last 15 trials was used in obtaining scores for this portion of extinction. The *t*-test is based on the distribution of the second-order differences indicated in the table, over the nine replications of the experiment.

Table IV shows that the difference in resistance to extinction between 100 percent and partial reinforcement is significantly greater with massed training than with spaced training. Results in the same direction were obtained when the analysis shown in Table IV was performed separately for those groups whose extinction trials were massed and for those groups whose extinction trials were spaced, but the differences were not nearly as reliable as in the overall analysis. For massed extinction analyzed separately, the probabilities were .07 for all extinction trials and .12 for the last half of the extinction trials that the difference obtained would arise by chance. For spaced extinction the corresponding probabilities were .20 and .11.

Discussion

As stated in greater detail in the introduction, the present experiment was based on a hypothesis which provides a stimulus-response explanation for the greater resistance to extinction that has been found after training with partial reinforcement as compared with 100 percent reinforcement. According to this explanation, the partial reinforcement technique has the effect of conditioning the response to cues normally present only during extinction. After one or more nonreinforced trials, the conditioned stimulus pattern at the start of a reinforced trial includes after-effects of the preceding nonreinforced trials, which make it much like a trial during extinction. When performance of the response in the presence of these nonreinforcement or extinction cues is now reinforced in the trial being considered, the partial reinforcement group learns to give the response to cues for extinction. When the actual extinction trials are started, the omission of reinforcement does not introduce unfamiliar cues; the response has been conditioned to nonreinforcement cues during training and continues at a relatively high

TABLE IV

Comparison of the Differences * Between 50% and 100% Reinforcement Groups During Extinction After Massed Training and After Spaced Training

	Massed Training			Spaced Training			D_1 (massed) minus D_2 (spaced)	$\sigma_{D_1-D_2}$	t	df	P**
	Mean Score		Mean Diff.	Mean Score		Mean Diff.					
	50%	100%		50%	100%						
All 30 ext. trials (Md. = 16.3 sec.)	17.8	14.2	3.6	13.6	14.4	−0.8	4.4	1.95	2.3	8	.03
Last 15 ext. trials (Md. = 25.3 sec.)	9.1	6.3	2.8	6.8	7.8	−1.0	3.8	1.09	3.5	8	<.01

* The measure used was mean number of responses at or below the overall median response time.
** For one tail of distribution.

level. After training with 100 percent reinforcement, on the other hand, the start of extinction introduces nonreinforcement cues for the first time. The conditioned response to the changed stimulus pattern is therefore weakened because it is at a different point on the generalization gradient from the reinforced stimulus pattern.

From this theoretical analysis it was predicted that spacing of training trials with animal subjects should reduce or destroy the advantage of partial reinforcement during extinction, because with long inter-trial intervals the after-effects of reinforcement or nonreinforcement on the preceding trials would have dissipated to some extent. The conditioned stimulus pattern at the start of a new trial would thus be more similar for the partial and 100-percent reinforcement groups during spaced training, and their performance during extinction should be more comparable. The results of the experiment bear out the prediction. Whereas after massed training, significantly greater resistance to extinction was found for 50 percent than for 100 percent reinforcement, after spaced training, there was no advantage for 50 percent reinforcement—the slight difference obtained was in the reverse direction.

The reversal obtained with spacing in the present results is not a necessary implication of the hypothesis, which implies only that the difference will be reduced. By changing the spaced intervals used, the degree of overlearning (i.e., amount of training with nonreinforcement followed by reinforcement), and other details of the procedure, it may be possible to vary the size of the difference from a small change, with 50 percent reinforcement still showing considerably greater resistance to extinction, to a significant reversal, with 50 percent reinforcement extinguishing more rapidly. Another factor affecting the size of the difference, and one which is probably little influenced by spacing of trials, is *reinstatement* of cues for continued responding. There are undoubtedly stimuli from the animal's behavior which are reinstated each time the animal is placed in the apparatus and which become cues for continued responding in spite of nonreinforcement. For example, in addition to cues from the frustration which carries over from a nonreinforced trial to the start of the following trial with massed training, there may be cues from *conditioned* frustration aroused by apparatus cues on each trial with *both* massed and spaced training. If the latter type of cue is a big part of the stimulus pattern, spacing of trials would not reduce by much the difference between extinction following 50 percent and 100 percent reinforcement.

As stated above, the hypothesis is applied specifically to a comparison of extinction after partial and 100 percent reinforcement. It may be applied more generally as a hypothesis to explain part of the decrement that characterizes all extinction. According to this hypothesis, a response becomes conditioned during training to the after-effects of reinforcement as part of the conditioned stimulus pattern. Omission of reinforcement during extinction thus changes the conditioned stimulus pattern, not only by the removal of reinforcement cues, but also by introducing whatever new cues result from the omission, such as those that are produced by the subject's reaction to the omission. Part of the decrement in the response during extinction would thus be interpreted as a weakening of the

response through generalization to the new stimulus pattern, combined with the weakening effects of any incompatible responses that might be produced by the new stimulus pattern or the omission of reinforcement.

It might be asked whether this hypothesis (as applied to partial reinforcement) can explain cases in which experimenters failed to obtain greater resistance to extinction with partial reinforcement. Humphreys (**14**), in a bar-pressing experiment with rats, found an advantage for partial reinforcement only if number of reinforcements was held constant. Four groups received reinforcements on (1) 18 out of 52 trials, (2) 18 out of 18 trials, (3) 7 out of 18 trials, and (4) 7 out of 7 trials. Comparison of number of responses during fixed extinction periods showed an advantage for (1) over (2) and an advantage for (3) over (4), in each case the number of reinforcements being the same and the number of trials variable. But there was no difference between (2) and (3), where number of reinforcements is variable and number of trials is constant.

The explanation of the effects of partial reinforcement proposed in the present report assumes that there must be a large enough number of occurrences of nonreinforcement followed by reinforcement during training for a stable response to be conditioned to the extinction cues. In (3) above, only seven reinforcements were given. If it is assumed that the first trial was always reinforced, there would be a maximum of six reinforced trials following nonreinforced trials. It is unlikely that a very strong response could be conditioned to the extinction cues in so few trials. If there had been a control group for (1), with 52 trials all reinforced, it is possible that a comparison of this group with (1) might have revealed a difference in resistance to extinction.

Another experiment which failed to demonstrate any difference in extinction between partial and 100 percent reinforcement is one by Denny (**4**), involving rats in a T-maze. However, this failure actually provides confirmation for the results of the present study, for the absence of an effect was under circumstances in which the present findings indicate there should be no effect, namely following training with widely spaced trials. Partial reinforcement groups were given one reinforced and one nonreinforced trial per day, and 100-percent reinforcement groups were given two reinforced trials per day (plus, in each case, two trials to the incorrect side, the last two trials of the day being forced to achieve this combination). But the interval between trials given on the same day was 20–30 min., and 24 hours intervened between every two correct choices. In the present study, it was shown that a 15-min. interval was sufficient to reduce to zero the difference in extinction of a running response following partial and 100 percent reinforcement.

Finger (**5, 6**), on the basis of his studies, concluded that partial reinforcement per se did not result in greater resistance to extinction, but his results are shown to be inconclusive in a study by Lawrence and Miller (**15**) because of the effects of certain aspects of Finger's technique.

The question arises as to whether other theories that have been proposed to explain the results of partial reinforcement are able to account for the results of the present study. Foremost among other explanations is the 'expectancy' interpretation proposed by Humphreys (**11**) and later amplified by Hilgard (**8**). A difficulty in this case is that the interpretation is not stated rigorously enough for its application in a new situation to be clear. In fact, its application even to Humphreys' findings is not unambiguous. The relatively higher level of responding during extinction after partial reinforcement is explained merely as due to the difficulty of shifting from an expectation of intermittent reinforcement to an expectation of complete nonreinforcement, compared with the ease (after 100 percent reinforcement) of shifting from an expectation of uniform reinforcement to an expectation of uniform nonreinforce-

ment. One might as readily argue, at the same level of discourse, just the reverse: that after being habituated to infrequent reinforcements, the subject finds it easy to get used to the idea that there will be none at all, but after being habituated to uniform reinforcement, the subject finds it hard to believe that there will be no more reinforcements.

The point is that Humphreys' explanation did not follow from an expectancy principle; it was an ex post facto conjecture as to what the course of expectancies must have been if expectancies are to be considered the causes of the behavior. The fact that in a subsequent study (12) Humphreys showed that the course of verbal expectancies did roughly correspond to that of conditioned responses is treated by Hilgard (8) as confirmation of this conjecture and as evidence in favor of an expectancy principle. However, Humphreys' subsequent results show only that verbal responses are affected by partial reinforcement in much the same way as other responses. And the critical issue is that neither of these was deduced from the expectancy hypothesis; the course of expectancies was determined empirically, and the course of the other conditioned responses was explained as being due to the course of expectancies. One of the advantages of the present hypothesis is that it can predict the course of both kinds of responses without recourse to any empirical analysis.

Humphreys' (11, 13) initial rise in the extinction curve after partial reinforcement is attributed to high expectation of reinforcement after two extinction trials because there never were more than two successive nonreinforcements during training. This must mean that the subject is keeping track of (i.e., responding differentially to the after-effects of) the successions of reinforcements and nonreinforcements. To say that, after a series of events which has always been followed by reinforcement, the subject has a high expectation of reinforcement contributes nothing beyond the stimulus-response interpretation that to a differential cue that has always been followed by reinforcement, there is maximal probability of eliciting the conditioned response.

If expectancy is viewed as the basic principle for explaining all learning, then there would appear to be no reason for predicting anything different for resistance to extinction following spaced training from that following massed training. The fact that the animals in this study *learned* as well, if not better, with spaced as compared with massed training would indicate that the expectancies used by the expectancy principle can be set up and maintained through intervals as long as those used. Therefore, it is difficult to see why spacing of training should remove the advantage in resistance to extinction usually found with partial reinforcement. Presumably expectancies are aroused when a previously-experienced situation recurs, regardless of the time interval separating the two exposures. Distributed practice should facilitate the acquisition of an expectancy as much as it does the learning of any response, and the advantage of partial reinforcement should be increased, if anything, by the spaced practice in forming expectancies.

However, if expectancies are regarded as secondary phenomena sometimes involved in learning, then such expectancies, particularly in the form of implicit verbal responses with human subjects, would be included here as one of the after-effects of reinforcement or nonreinforcement as described in the present hypothesis. For human subjects, such verbal responses would undoubtedly be an important part of the reinforcement or extinction cues to which the response becomes conditioned. For example, human subjects who tended to verbalize what went on in Humphreys' experiments (11, 13) might, after a certain amount of practice, think to themselves the equivalents of: "Two shocks in a row, none next time," or "No shock and then shock, maybe shock again this time," or "No shock for two trials, shock for sure next time." Such verbalizations would function as after-effects and would

provide distinctive cues exactly correlated with preceding patterns of reinforcement (to the extent that they correctly described the actual events of the experiment). They would also very probably be one of the cues reinstated at the start of each trial. Furthermore, they would be differential cues that had been reinforced, respectively, 0 percent, 50 percent, and 100 percent of the time in Humphreys' procedure. It is not surprising, therefore, that strength of actual responding correlated to some extent with cues so differentially reinforced. A separate expectancy principle is not required to explain why these expectancies help produce the greater resistance to extinction found by Humphreys. The subject who has always been reinforced while thinking "No shock for two times" is more likely to give the conditioned response when he thinks "No shock for three or four times" than the subject who not only has never been reinforced while thinking this but also never thought anything like it during acquisition. Such verbalizations function exactly as the presence or absence of food taste in the mouth and the state of frustration in the case of rats.

But applying the expectancy principle to animals as if they had implicit verbal cues would very probably lead to errors of prediction. Humphreys (14) himself does not attempt to apply the expectancy principle in his study of partial reinforcement with rats. In the case of human subjects, spacing of training trials would probably have much less effect of reducing the difference between partial and 100 percent reinforcement in resistance to extinction, because verbal cues could doubtless still be re-aroused after long inter-trial intervals. Still, some effect similar to the present findings should show up since the advantage of partial reinforcement would partly depend, in human subjects, on their ability to remember, at any point during learning, what had happened on the last few trials, and since some forgetting would be expected with long intervals between trials.

Another proposal for explaining the effects of partial reinforcement is one by J. S. Brown, which has been tested by Mowrer and Jones (17). They have designated it the 'response-unit' hypothesis. Speaking of their bar-pressing response, they suggest that what is reinforced in the partial reinforcement situation is not the single bar-depression that immediately precedes food but the sequence of bar-depressions leading up to the reward. Defining the 'response' as a sequence or pattern of behavior rather than a single act, they propose that the behavior temporally more remote from the reward is reinforced by it to a diminishing extent according to the principle of the gradient of reinforcement. Accordingly, animals rewarded during training for several depressions of the bar would be expected during extinction to press the bar more times but not necessarily to make more 'responses' than animals whose 'response' during training was a single bar-depression.

From this hypothesis it would probably be predicted that spacing of training trials would make it difficult for a sequence of behavior to be reinforced as a 'unit' and would therefore cause a breakdown in the usual advantage of partial reinforcement in extinction. However, even with massed training, it would appear to be much easier to conceive of several bar-depressions as a response-unit than two or three runs down an alley separated by 15-second intervals in the home cage. In the latter case, the length of time involved and the possibility for variable intervening activity make the idea of a response-unit untenable.

Denny (4) has obtained evidence supporting a secondary reinforcement hypothesis for explaining the fact that *acquisition* is almost as good with partial reinforcement as with 100 percent reinforcement. He does not attempt to apply it to differences in *extinction* under the two conditions, which in fact his study does not show. Secondary reinforcement from the empty goal box unquestionably operated in the 50 percent groups in the present study, but there is

no reason to expect it to operate differently when training trials are spaced from when they are massed. Secondary reinforcement not only cannot explain the differential effect on resistance to extinction obtained with massed and spaced training trials, but also it cannot account for the basic difference in extinction between partial and 100 percent reinforcement with massed training trials. In order for the secondary reinforcement principle to explain the latter, it would have to be assumed that secondary reinforcement is stronger than primary reinforcement.

Summary

1. Seventy-two rats were trained to run down an alley for food. Half received reinforcements on all training trials, and half randomly on 50 percent of the trials. Half of each group were trained with a 15-sec. interval between trials and half with a 15-min. interval. Each of the four training groups was divided for extinction, half being extinguished with the 15-sec. interval and half with the 15-min. interval.

2. There were no significant differences in level of performance on the last half of the acquisition trials, either between 100 percent and 50 percent reinforcement or between massed and spaced training.

3. After massed training, resistance to extinction was significantly greater for 50-percent reinforcement groups than for 100-percent reinforcement groups.

4. After spaced training, the difference in resistance to extinction between 100-percent and 50-percent reinforcement groups was not signiffcant; it was, in fact, slightly reversed.

5. The differential effect of partial reinforcement depending on whether training was massed or spaced was found to be significant.

6. These results verify a prediction from a hypothesis utilizing stimulus-response learning concepts for explaining the effect of partial reinforcement on extinction.

(Manuscript received November 5, 1948)

References

1. Brogden, W. J. The effect of frequency of reinforcement upon the level of conditioning. *J. exp. Psychol.*, 1939, **24**, 419–431.
2. Brunswik, E. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, **25**, 175–197.
3. Cole, L. E. A comparison of the factors of practice and knowledge of experimental procedure in conditioning the eyelid response of human subjects. *J. gen. Psychol.*, 1939, **20**, 349–373.
4. Denny, M. R. The role of secondary reinforcement in a partial reinforcement learning situation. *J. exp. Psychol.*, 1946, **36**, 373–389.
5. Finger, F. W. The effect of varying conditions of reinforcement upon a simple running response. *J. exp. Psychol.*, 1942, **30**, 53–68.
6. Finger, F. W. Retention and subsequent extinction of a simple running response following varying conditions of reinforcement. *J. exp. Psychol.*, 1942, **31**, 120–133.
7. Guthrie, E. R. *Psychology of learning.* New York: Harper, 1935.
8. Hilgard, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.
9. Hull, C. L. Notes on the Humphreys extinction paradox. Unpublished symposium notes, Midwestern Psychol. Assn., 1941.
10. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
11. Humphreys, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.
12. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.
13. Humphreys, L. G. Extinction of conditioned psychogalvanic responses following two conditions of reinforcement. *J. exp. Psychol.*, 1940, **27**, 71–75.

14. HUMPHREYS, L. G. The strength of a Thorndikian response as a function of the number of practice trials. *J. comp. Psychol.*, 1943, **35**, 101–110.
15. LAWRENCE, D. H., & MILLER, N. E. A positive relationship between reinforcement and resistance to extinction produced by removing a source of confusion from a technique that had produced opposite results. *J. exp. Psychol.*, 1947, **37**, 494–509.
16. MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
17. MOWRER, O. H., & JONES, HELEN. Habit strength as a function of the pattern of reinforcement. *J. exp. Psychol.*, 1945, **35**, 293–311.
18. PAVLOV, I. P. *Conditioned reflexes.* (Trans. by G. V. Anrep) London: Oxford Univ. Press, 1927.
19. SHEFFIELD, VIRGINIA F. Resistance to extinction as a function of the distribution of extinction trials. *J. exp. Psychol.*, 1950, **40**. (In press)
20. SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
21. TIPPETT, L. H. C. *Random sampling numbers.* London: Cambridge Univ. Press, 1927.

DISTRIBUTED PRACTICE IN VERBAL LEARNING AND THE MATURATION HYPOTHESIS *

BY SUSAN T. H. WRIGHT AND DONALD W. TAYLOR

Stanford University

INTRODUCTION

The hypothesis of stimulation-induced maturation was first made explicit by Doré and Hilgard (1). Wheeler (5) had earlier advanced the idea that: "Repetition of stimulation is unquestionably a condition of learning, not because it leads to the use of the same nerve patterns over and over again but because it induces maturation when the stimulation is properly timed." Developing this line of thought, Doré and Hilgard (1) formulated the hypothesis that: "At any given age-level of maturation, distributed stimulation in a specific function will give rise to an improvement in that function at a rate which is related more closely to the time elapsing from the beginning of practice than with the total amount of practice."

The previous experiments relevant to this hypothesis have been reviewed by Hilgard (2) in his recent book, *Theories of learning*, and with one exception will not be described here. In a chapter devoted to Wheeler's learning theory, Hilgard uses these experiments to illustrate the emphasis placed on maturation by organismic psychology.

The previous experiments all involved perceptual-motor learning. The purpose of the present study was to test the applicability of the hypothesis to rote verbal learning.

The experimental design employed in the present investigation was deliberately patterned after that used in Doré and Hilgard's original study (1). In that experiment, four groups of 25 *Ss* practiced individually on the Koerth pursuit rotor during a single 43-min. period. The individuals in each group were given three initial one-min. trials separated by one-min. rests; all groups were also given a final one-min. test trial following a one-min. rest period. Following the initial trials, three of the groups practiced for one-min. periods and rested 11, 3 and 1 min. between trials, respectively. These groups received a total of 7, 13 and 22 trials, respectively, during the 43-min. period. The fourth group received three-min. practice periods alternated with one-min. rest periods for a total of 31 trials.

The latter three of the four groups made essentially the same mean score during the 41st min., this in spite of the fact that out of the 41 min. they had spent 12, 21 and 30 min., respectively, in practice. These results are consistent with the hypothesis of stimulation-induced maturation. The learning was proportional to elapsed time rather than to the amount of practice. The achievement in 41 min. of the 7-trial group was definitely inferior to that of the other groups. In terms of the maturation hypothesis, it would be said that the stimulation provided the 7-trial group was insufficient for optimum growth, that the stimulation for the 13-trial group was somewhere near optimum, while that for the 22-trial and 31-trial groups was excessive.

* This article is based on experimental work incorporated in a thesis submitted by Miss Wright to Stanford University in partial fulfillment of the requirements for the degree of Master of Arts. The second author supervised the experimental work and prepared the present report.

It should be added that, although their experimental results are consistent with it, Doré and Hilgard (1) express reservations with respect to the maturation hypothesis. They point particularly to fatigue and to the tendency of massing of practice to conceal the effects of learning as two factors which may provide at least a partial alternative explanation.

The Experiment

The task set the *Ss* in the present study was that of learning a list of 18 nonsense syllables presented serially at the rate of one every three sec. by means of a Cornell Memory Drum. The method of serial anticipation was employed. Before the first syllable was a cue symbol, o–o–o, and after the last syllable a blank space. Thus, for each trial there were 20 exposures of three sec. each and each trial lasted 1 min.

A total of 80 *Ss* were obtained from the elementary class in psychology. These were divided into five groups of 16 *Ss* each on the basis of their performance on the second preliminary trial. In this way the mean scores of the five groups on the second trial were equated.

All five groups were given a total practice period of 52 min. Following the initial presentation of the list, all groups were given two preliminary trials with two min. rest intervals interposed. Beyond this point, the spacing of practice was systematically varied among the several groups. Group A received eight-min. rest periods. Groups B, C, and D were given rest periods of 3¼ min., 2 min., and 1¼ min., respectively. Group E practiced continuously. During the rest intervals *S* was given a cartoon book to look at. This schedule was designed so that each group would have a trial during the 16th, 25th, 34th, 43rd, and 52nd min. of practice. The purpose of this feature of the design was to facilitate comparisons of group achievement.

Two lists of 18 nonsense syllables each were taken from the study of Melton and Irwin (3). Half of the *Ss* in each group learned one list; the other half learned the other list. The *Ss* were instructed to spell the syllables, not to pronounce them. If an *S* learned all 18 syllables to a criterion of two successive perfect trials, he was allowed to stop. Eleven of the 80 *Ss* met the criterion in less than the scheduled number of trials. Their data are included as though they had continued to score correctly on the remaining trials. However, no *S* had reached the criterion by the 12th trial and only one had done so by the 34th min., the two points at which the tests of significance of differences described below were carried out.

Results

The results obtained are of interest first with respect to the general problem of optimum spacing of practice. Fig. 1 presents the mean achievement of each of the five groups as a function of the number of trials of practice. The data show clearly that spaced practice is superior to massed practice. But the more interesting finding is that varying the length of the interpolated rest periods from 1¼ min. to 8 min. produced no significant variation in the achievement per trial of practice. Fig. 1 shows that the learning curves for the four groups with distributed practice cross and recross each other repeatedly. At no point is there a significant difference between the mean scores of any two of these groups.

In contrast, the learning curve for Group E with continuous practice is consistently inferior to those of the other four groups. The significance of this difference may be tested at the 12th trial, this point being chosen for two reasons: It was the last trial for Group B and, secondly, no *S* at this point had met the criterion. On the 12th trial the means of Groups B, C, and D were 9.63, 9.38 and 9.44 respectively. Since there were no significant differences among them, the three groups were combined into one with a mean of 9.48. The difference between this mean and that of 7.19 for Group E was tested by means of the *t*-test for uncorrelated measures and found to be significant at the 5 percent level.

The finding that learning efficiency did not increase when the length of the interpolated rest periods was increased from 1¼ min. to 8 min. contrasts with the results obtained in two

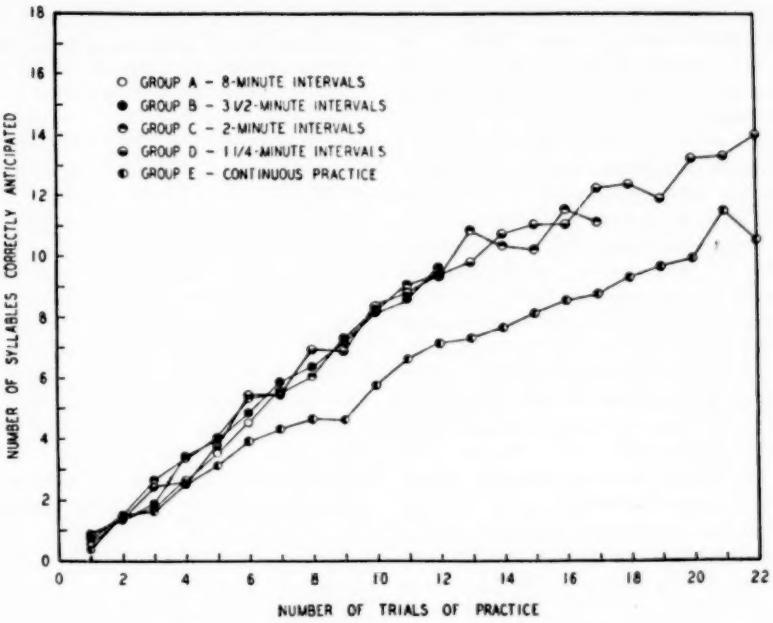

FIG. 1. Achievement as a function of amount of practice

other studies involving the same range of length of rest periods. Thus, in the study described above, Doré and Hilgard (1) found that increasing the rest periods from 1 min. to 11 min. did increase the efficiency of learning on the Koerth pursuit rotor. The discrepancy between their finding and that of the present study may possibly be due to the difference in the type of learning task employed. Piéron's study (4), however, also involved rote verbal learning, the task being that of learning lists of 20 numerals to mastery. The rest periods employed included 30 sec., 2 min., 5 min., and 10 min. As the length of the rest periods was increased, the number of trials required for mastery decreased from 11 trials with 30-sec. intervals to 5 trials with 10-min. intervals. The explanation of this difference in results is not apparent, although it should be noted that Piéron's study involved only one subject.

The present data do not confirm the predictions made on the basis of the hypothesis of stimulation-induced maturation. Fig. 2 presents the achievement of the five groups as a function of elapsed time instead of number of trials of practice. The maturation hypothesis would lead one to expect that the curves for two or more of the groups would coincide. This expectation follows from the view that maximum achievement within a given period of time will be produced by a certain optimum amount of stimulation or practice, that additional practice within that period will produce no additional achievement. Suppose, for example, that the hypothesis were valid and that the 17

trials given Group C represented optimum stimulation. Then the larger number of trials given Groups D and E would have produced no additional achievement. The achievement of the three groups would have followed the same general curve when plotted as a function of elapsed time.

The data presented in Fig. 2 do not support the maturation hypothesis. Instead, the results show that the larger the number of trials given in 52 min. the greater the achievement, the limiting case being that of Group E, with continuous practice, whose performance was consistently superior to that of any of the groups with distributed practice.

The *t*-test for uncorrelated measures was used to determine the significance of the differences between the mean scores of adjacent groups during the 34th min. The 34th min. was chosen for two reasons: All groups had a trial at this point and, secondly, only one *S* had yet reached the criterion. By the 43rd min. eight *S*s, and by the 52nd min. a total of eleven, had reached the criterion. Table I presents the results of the

TABLE I

SIGNIFICANCE OF DIFFERENCES BETWEEN MEAN SCORES OF ADJACENT GROUPS DURING THE 34TH MIN. OF PRACTICE

Group	Mean Score (34th min.)	Difference between Adjacent Group Means	*t* ($df = 30$)	*P*
A	3.56			
		2.81	3.13	.01
B	6.37			
		2.75	2.33	.03
C	9.12			
		1.63	1.34	.20
D	10.75			
		2.69	2.40	.03
E	13.44			

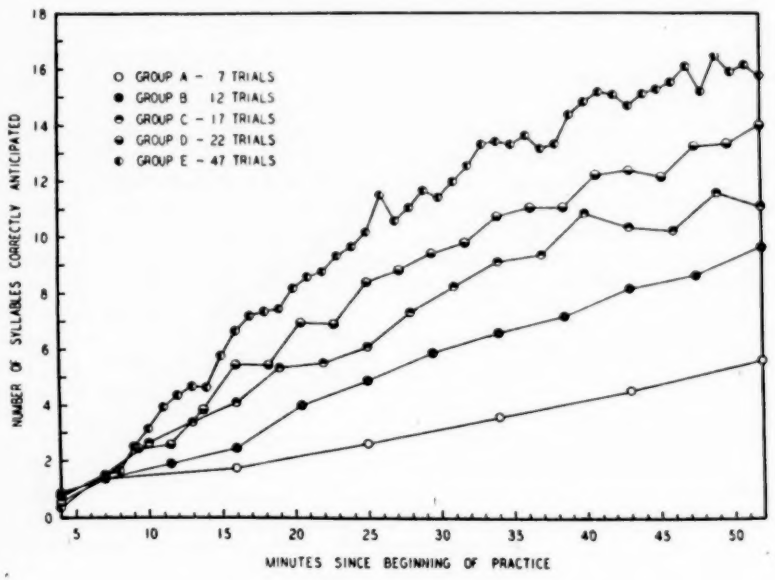

FIG. 2. Achievement as a function of elapsed time

analysis. The difference between adjacent means are significant in all but one case. Of primary interest is the fact that the difference between Group E with continuous practice and group D, the next best, is significant at the three percent level.

It may be noted that the test of significance employed did not take into account the fact that the five groups were equated with respect to their means on the second trial. Had this been done the levels of significance would undoubtedly have been higher. However, more elaborate statistical analysis does not seem necessary to justify the conclusion that the present data do not support the hypothesis of stimulation-induced maturation.

Summary

The purpose of the present study was to test the applicability of the hypothesis of stimulation-induced maturation to rote verbal learning.

Five groups of 16 *Ss* were employed. That task was that of learning by the method of serial anticipation a list of 18 nonsense syllables presented at the rate of one every three sec. All *Ss* received a total practice period of 52 min., and all trials were one min. in length. The length of the interpolated rest periods was systematically varied. One group of *Ss* practiced continuously. The other four groups received rest periods of 8, $3\frac{1}{2}$, 2, and $1\frac{1}{4}$ min. respectively.

The achievement per trial of all four groups with spaced practice was superior to that of the group with continuous practice. The more interesting finding is that varying the length of the interpolated rest periods from $1\frac{1}{4}$ to 8 min. produced no significant variation in achievement per trial.

The maturation hypothesis predicts that maximum achievement within a given period of time will be produced by a certain optimum amount of stimulation or practice, that additional practice within that period will produce no additional achievement. The results obtained do not confirm the hypothesis. Instead, the data show that the larger the number of trials given in 52 min. the greater the achievement. In terms of achievement per unit of elapsed time, the performance of the group with continuous practice was significantly better than of any of the groups with distributed practice.

(Manuscript received
September 24, 1948)

References

1. Doré, L. R., & Hilgard, E. R. Spaced practice and the maturation hypothesis. *J. Psychol.*, 1937, **4**, 245–259.
2. Hilgard, E. R. *Theories of learning.* New York: Appleton Century Crofts, Inc., 1948. Pp. 250–257.
3. Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173–203.
4. Piéron, H. Recherches expérimentales sur les phénomènes de mémoire. *Année psychol.*, 1913, **19**, 91–193.
5. Wheeler, R. H. *The science of psychology.* New York: Crowell, 1929, p. 299.

ABILITY PATTERNS UNDER DISTRIBUTED PRACTICE[1]

BY MARY J. KIENTZLE

The State College of Washington

In an earlier study (2), the writer found definite relationships between group variability and distribution of practice. Data from 12 groups of subjects who practiced printing the inverted alphabet under varying conditions of spacing, showed conclusively that:

1. Trial by trial, groups with spaced practice obtained higher mean scores than groups with massed practice.

2. Group variability rose more rapidly with spaced practice than with massed practice. The sharper rise in variability still appeared when size of mean scores was taken into account.

3. Correlations between scores on specific trials were associated with the ordinal numbers of the trials, but not with the amount of rest between them. In other words, intertrial correlations were invariant under spacing.

As it stood, the third conclusion was subject to varying interpretations. Correlations among scores on Trials 1, 2, 5, 10, 12, 15, 19, and 20 from the massed groups showed the same patterns, and were of the same size, as the corresponding correlations from the various spaced groups. Yet, there was no way to be sure that factors accounting for the correlations in one experimental group were the same as factors in a group which had practiced under different conditions. Even though the 12 groups were working under comparable conditions during their first trial, correlations of scores on Trial 1 with scores on subsequent trials were not high enough to say positively that similar correlation patterns for the diverse groups implied the operation of the same underlying factors for all experimental conditions. It was the kind of indeterminacy that would arise if the correlation between Tests X and Y were .60, and the correlation between Tests X and Z were also .60. Since the coefficient of alienation in each case would be .80, there would be no guarantee that any of the factors common to X and Y were also common to X and Z.

On the other hand, if r_{XY} were .95 and r_{XZ} were also .95, we could be fairly certain that X, Y, and Z were largely determined by the same factors. It has been found repeatedly that correlations among amount-done-per-unit-time scores from trials later in practice are much higher than correlations among scores from earlier trials. In fact, correlations among scores on later trials, when corrected for attenuation, often come close to 1.00. The use of trials later in practice as departure points for varying conditions of distribution should yield fairly unequivocal data regarding the similarity of learning processes under massed and distributed practice.

With this in mind, the present experiment was designed to show more clearly the relationship between group variability and the distribution of practice.

PROCEDURE

Four groups of *S*s practiced writing the inverted alphabet for 20 one-min. trials under the following experimental conditions:

[1] This research was supported in part by a grant-in-aid from the College Committee on Research of the State College of Washington for 1947–1948.

20M *Group A* ($N = 78$) had 20 trials with no rest at all between trials.

15M5D *Group B* ($N = 92$) had 15 trials with no interpolated rest. Trials 16, 17, 18, 19, and 20 were each preceded by 60 sec. of rest.

20D *Group C* ($N = 94$) had 20 trials with 60 sec. of rest between trials.

15D5M *Group D* ($N = 81$) had 15 trials with 60 sec. of rest between trials. Trials 16, 17, 18, 19, and 20 were each preceded by no rest at all.

The *S*s were elementary psychology students at the State College of Washington. They participated in the experiment during regular class periods. They were instructed to print the alphabet upside down as rapidly and as accurately as possible, making one letter in each square of the four-to-the-inch cross-section paper provided as test blanks. As soon as they completed a line, they covered it with a 'cover sheet' to preclude the possibility of copying from one line to the next. During rest periods, the *S*s were asked not to look at their papers, and not to think or to talk about the task. When no rest pause was interpolated, they worked immediately on the next trial.

The test blanks consisted of two 8½ × 11 sheets of four-to-the-inch cross-section paper, to which was attached the cover sheet mentioned above. The two record sheets were numbered along the right-hand side to indicate the starting point for each trial. For all groups, score was the number of letters printed per trial, regardless of accuracy.

Results

Table I shows the means and standard deviations at each trial for all

TABLE I

Means and Standard Deviations of Distributions at Each Trial

Condition	Group	N		Trial									
				1	2	3	4	5	6	7	8	9	10
20M	A	78	M	29.0	32.8	36.5	37.9	39.2	40.8	42.7	44.8	43.8	45.4
			σ	9.6	9.1	9.3	9.6	10.0	10.5	9.6	9.1	9.6	10.2
			σ/σ_1	1.00	.95	.97	1.00	1.04	1.09	1.00	.95	1.00	1.06
15M5D	B	92	M	28.4	29.7	33.7	36.1	38.2	38.6	40.7	42.0	42.3	43.1
			σ	10.3	9.4	9.5	10.7	10.3	10.6	10.9	11.1	10.2	10.4
			σ/σ_1	1.00	.91	.92	1.04	1.00	1.03	1.06	1.08	.99	1.01
20D	C	94	M	26.4	33.2	41.0	45.0	48.7	50.3	53.4	56.4	57.8	58.5
			σ	8.8	9.9	10.6	10.8	10.6	10.6	10.7	10.9	10.6	10.1
			σ/σ	1.00	1.12	1.20	1.23	1.20	1.20	1.22	1.24	1.20	1.15
15D5M	D	81	M	26.9	35.8	41.2	47.3	51.2	53.6	56.3	58.6	60.6	62.8
			σ	10.0	12.3	12.4	12.7	11.8	12.2	11.9	11.8	11.9	11.8
			σ/σ_1	1.00	1.23	1.24	1.27	1.18	1.22	1.19	1.18	1.19	1.18
				11	12	13	14	15	16	17	18	19	20
20M	A	78	M	43.8	47.8	48.9	49.3	49.0	49.4	50.8	52.4	52.6	55.3
			σ	10.4	9.7	10.5	10.6	10.5	11.4	11.5	10.6	10.4	11.8
			σ/σ_1	1.08	1.01	1.09	1.10	1.09	1.19	1.20	1.10	1.08	1.23
15M5D	B	92	M	43.3	44.8	46.6	46.6	47.0	54.7	58.0	59.3	61.4	62.0
			σ	10.7	11.1	11.2	11.0	10.9	12.8	13.0	12.8	12.8	13.1
			σ/σ_1	1.04	1.08	1.09	1.07	1.06	1.24	1.26	1.24	1.29	1.27
20D	C	94	M	57.5	62.2	62.8	64.3	65.9	65.0	66.7	68.5	70.0	71.8
			σ	10.9	11.2	11.4	11.3	11.4	11.4	12.1	11.3	11.4	11.4
			σ/σ_1	1.24	1.27	1.30	1.28	1.30	1.30	1.37	1.28	1.30	1.30
15D5M	D	81	M	62.3	65.7	66.7	69.4	68.3	60.4	62.7	58.4	59.7	62.2
			σ	12.2	12.9	12.6	12.4	12.9	10.9	10.9	10.1	11.0	11.3
			σ/σ_1	1.22	1.29	1.26	1.24	1.29	1.09	1.09	1.01	1.10	1.13

four groups. Also, it shows the variability at each trial as a proportion of the initial variability of the particular experimental group.

The data in this table support the conclusions of the earlier study that distributed practice yields higher mean scores than massed practice. Furthermore, they agree with Gentry's (1) conclusions that changing from massed to spaced practice and vice versa changes mean scores rapidly. Group B, which changed from massed to spaced practice, showed an immediate increase in mean score with Trial 16, and continued to improve through the 20th trial. Group D, which changed from spaced to massed practice, showed lower mean scores at once. The last five mean scores earned by this group were all lower than the same group's mean score on Trial 15.

Also, when the ratio of the variability at any trial to initial variability is considered, it is apparent that variability rises much more slowly under massed practice. This fact is especially clear for Group D, which changed from spaced to massed practice. For Trials 16 through 20, the mean scores ranged from 58.4 to 62.7; the variabilities for these massed trials averaged eight percent higher than the initial variability. However, on Trials 8 through 11, when the means showed substantially the same range, the variabilities averaged 19 percent higher than the initial variability.

As in the earlier study (2), the slower rise of group variability with massed practice is not entirely contingent upon the lower mean scores earned by the massed groups. Rather when the size of the mean score is taken into account, the slower increase of variability for the massed groups is

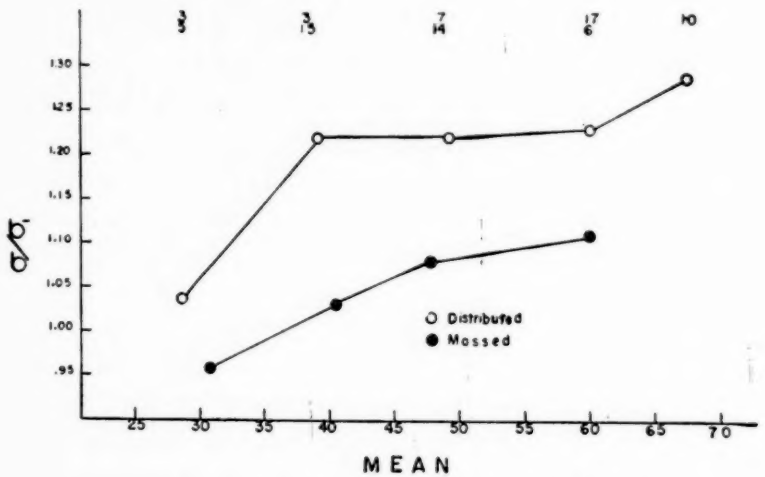

FIG. 1. Variability as a function of mean score and condition of distribution

Means obtained under massed conditions were grouped according to size (e.g., from 35 to 44) and their average was entered on the abscissa. The average of the σ/σ_1's corresponding to them were plotted as the ordinate. Points for distributed practice were similarly obtained. The first row of figures above the graph shows the number of observations entering each point on the upper curve; the second row gives the same information for the lower curve.

still apparent. Fig. 1 shows the relationship between the mean score and the ratio of the corresponding variability to the initial variability. Although variability rises with both conditions of practice, its progress is more rapid with interpolated rest.

Scores from Trial 15 were correlated with scores from Trial 16 and Trial 20. In order to estimate the reliability of scores on Trial 15, 16, and 20, individual practice curves for Groups A and C were plotted and smoothed free-hand. The correlation between the observed and smoothed scores for a particular trial was considered its reliability. The reliabilities thus found ranged from .91 to .97 and showed no systematic variation with either ordinal number of the trial or with experimental condition. Since differences among these six coefficients were insignificant ($r = .95$), the value corresponding to the average of the coefficients transformed by Fisher's z-method was used as the reliability of the scores on all three trials for both groups. The reliability of scores on these three trials for Groups B and D was also assumed to be .95.

The correlations, both observed and corrected for attenuation, are presented in Table II.

Fisher's z-test showed differences among the raw correlations $r_{15,16}$ to be no larger than might be expected from errors of random sampling. Similarly, $r_{15,20}$ from the four groups also showed only chance variation. That is, there were no significant differences among correlations for the four experimental groups.

TABLE II

CORRELATIONS BETWEEN SCORES ON TRIAL 15 AND SCORES ON LATER TRIALS

Condition	Group	N	Observed Coefficients		Corrected Coefficients	
			$r_{15,16}$	$r_{15,20}$	$r_{15,16}$	$r_{15,20}$
20M	A	78	.853	.901	.900	.951
15M5D	B	92	.901	.902	.951	.952
20D	C	94	.918	.880	.969	.929
15D5M	D	81	.847	.848	.894	.895

TABLE III

SIGNIFICANCE (P) OF DEVIATIONS OF OBSERVED SCORES FROM NORMALITY

Condition	Group	Trial 15	Trial 16	Trial 20
20M	A	.78	.54	.66
15M5D	B	.99	.96	.76
20D	C	.66	.88	.93
15D5M	D	.65	.33	.54

Since correlations may be affected by the shape of the distributions from which they come, observed scores on Trials 15, 16, and 20 were fitted to normal curves. Table III shows the significance of deviations from normality, as found by the chi-square technique.

It is evident from Table III that deviations from normality do not require the rejection of the null hypothesis.

DISCUSSION

The results described herein imply that rest and practice have different functions in improving scores.

The increased variability under distribution, which appeared not only in the present study, but also in an earlier one employing a wide variety of rest intervals (2), indicates that *Ss* varied widely in processes occurring during rest intervals. Although further work would be required to determine exactly the relative contributions of practice and rest to the extent of group variability, it is evident that rest was associated with an increase in variability. Furthermore, the increased variability with interpolated rest was not a mere accompaniment

of the higher mean scores earned by the rested groups. When mean scores of the same size were considered, higher variabilities accompanied the means obtained under spaced conditions.

Although changing conditions of spacing late in practice materially affected mean scores and standard deviations, it did not affect correlations with the 15th trial. Gains between Trial 15 and later trials varied, but correlations with Trial 15 did not.

Looking upon Group A as a control for Group B, we see that although Group A gained only 6.3 points between Trials 15 and 20, Group B, which shifted to spaced practice, gained 15 points. Nevertheless, $r_{15,20}$ is the same for both groups. Since, for the first 15 trials, Group A and Group B both used massed practice, the explanation of the variance of Trial 15 in the two groups should logically involve the same abilities. Similarly, the explanation of the variance of Trial 15 should be the same for Group C as for Group D.

In Group B, $.95^2$, or 90 percent of the variance of Trial 16 was associated with variation in Trial 15; in Group A, $.90^2$, or 81 percent, of the variance of Trial 16 was associated with variation in Trial 15. In other words, the variance of Trial 16 from Group A and the variance of the same trial from Group B were both largely accounted for by variation in Trial 15 —which logically should have the same explanation in both groups. The same situation holds in considering the determination of Trial 20 from Groups A and B. And it holds in considering the determination of Trials 16 and 20 in Groups C and D. Neither in instances where the same condition of distribution continued nor in instances where the conditions were changed were the variances of later trials wholly determined from the variance of Trial 15. Nevertheless, correlations with scores on Trial 15 stayed high enough in Groups A and B to indicate substantial overlapping, if not identity, of determiners of Trial 16, and also of Trial 20, in the two groups. Likewise, the determiners of Trial 16, and of Trial 20 from Groups C and D, show considerable overlapping.

Although the above conclusion applies strictly only to the similarity of underlying abilities within the first two groups, and within the last two groups, the similarity of intertrial correlations for all four groups, along with the constancy of correlations under altered procedures, would lead one to suspect a similarity throughout of determiners for any particular trial. Furthermore, one might reasonably suppose that invariance of intertrial correlation matrices under varied distributions of practice (2) implies an actual invariance of abilities entering the task.

If further research confirms the suspected state of affairs, considerable light would be shed on the nature of practice. Abilities are variables postulated to account for individual differences in scores when experimental conditions are held constant. The rationale behind such a postulate comes from the assumption that when two sets of events vary together, they can be explained in terms of the same underlying variables. In the field of individual differences, we measure extent of association between two sets of scores by the correlation coefficient, and postulate independent abilities to account for that association. Thus, if $r_{15,16}$ is .95, we assume that scores on both trials reflect largely the same abilities.

If intertrial correlations depend on number of trials, but not on the

amount of intervening rest, and if the invariance of correlations means an invariance of component abilities, then rest does not change a subject's standard score on a specified trial. That is, if it were possible to observe, say, two 15th trials from the same group of subjects, the one trial under massed conditions and the other under spaced, each subject's standard score would be the same on both trials. He might earn a higher raw score with rest, but he would keep the same standing on both trials. In other words, the correlation between scores on the two hypothetical 15th trials would be 1.00. This in turn would imply that raw scores on the first hypothetical Trial 15 (X_{A15-1}) would be a linear transformation of raw scores on the second hypothetical Trial 15 (X_{A15-2}). Such a linear transformation might be written as follows:

$$X_{A15-1} = m\, X_{A15-2} + b,$$

in which m and b are constants. In order for $r_{(A15-1),(A15-2)}$ to remain 1.00, m and b would have to stay the same for all subjects. They would reflect the relationship between the means and standard deviations of scores on the two hypothetical trials.

Such a relationship between raw scores on the two trials would imply that the ability pattern of one trial is exactly the same as the ability pattern of the other trial, when 'scores' on the two abilities are standard scores. In other words, the relative contributions of various abilities to variance at any particular trial would be independent of rest.

The above conclusions are necessarily tentative. Nevertheless, the data of the present study, as well as those of an earlier one (**2**), suggest strongly that, while practice changes the underlying abilities entering a task, rest improves performance in some total fashion which leaves ability relationships unchanged. Level (as shown by mean scores) and variability of performance depend on rest as well as on practice, but ability pattern does not. A final answer to the question might be obtained by factor-analyzing scores from different stages of practice under varying conditions of distribution along with several tests of known factor content. The present investigation indicates that this or some similar analysis of individual differences would contribute materially to our understanding of the nature of learning.

(Manuscript received October 12, 1948)

References

1. Gentry, J. R. Immediate effects of interpolated rest periods on learning. *Teach. Coll. Contr. Educ.*, 1940, No. 799.
2. Kientzle, M. J. Properties of learning curves under varied distributions of practice. *J. exp. Psychol.*, 1946, **36**, 187–211.

THE PERCEPTION OF THE VERTICAL: I. VISUAL AND NON-LABYRINTHINE CUES *

BY CECIL W. MANN, NEWELL H. BERTHELOT-BERRY, AND HENRY J. DAUTERIVE, JR.

Tulane University

The orientation of the pilot flying in three dimensions of space presents many problems. It has been pointed out by Vinacke (13) that pilots with many hours of flying experience are subject to certain visual and postural errors of judgment which tend to reduce their flying efficiency. Adjustment to the flying situation brings into operation receptors in the semicircular canals, the otoliths, and the proprioceptive indicators in viscera, muscles, tendons, and the skin senses. The intact organism will use any or all cues which are appropriate to the situation. In broad daylight and with good visibility visual cues will probably predominate. Even under these conditions the importance of proprioceptive cues as an orientation factor cannot be neglected. Under conditions of reduced visibility—night flying, flying in clouds and overcast—proprioceptive cues will assume increased importance.

Many attempts have been made to establish the relationships existing between vestibular and visual functions in the maintenance of orientation. An adequate summary of this work up to 1934 has been given by Dusser de Barenne (5). It is significant that with changes in experimental design emphasis has been thrown at one time upon visual cues, at another upon vestibular and non-vestibular cues. In 1874–5, Mach, Breuer, and Crum Brown published independently their theories concerning the relation of the semicircular canals to postural and rotational perception. In 1875, Breuer differentiated the function of the receptors in the ampullae of the semi-circular canals and the maculae of the utricle and the saccule. Bourdon (4), from the evidence he obtained by varying the somesthetic cues and the position of the head with respect to the body, believed that the proprioceptive cues from the trunk and limbs were as important in postural judgment as were the vestibular functions. Garten (6) placed subjects in a 'nodding chair,' in air and in water, and from a comparison of their judgments, made under both conditions, concluded that 'muscle sensibility' was the predominant factor in postural judgments. Arndts (1) followed this by an experiment in which he used local anaesthetics to reduce the skin sensitivity, and concluded that 'deep sensibility' was sufficient for coarse postural orientation. Working with subjects who had lost vestibular function, Birren (3) claimed that his results showed that postural control may exist independently of visual and vestibular cues.

In a survey of the experimental work of the determinants of the perceived vertical and horizontal, Gibson and Mowrer (7) set up the hypothesis: "That both the visual and the postural vertical are determined by visual factors and gravitational factors acting jointly, with orientation to gravity, however, as the more decisive factor in cases of real conflict between the two types of sensory data and the primary factor genetically."

* This report is part of a series of investigations conducted jointly with the School of Aviation Medicine and Research, under contract N7onr-434, Task Order I, with the Office of Naval Research in cooperation with the Bureau of Medicine and Surgery and The Tulane University of Louisiana. Project Designation Number NR140-455 of the Medical Sciences Division, Office of Naval Research. Project No. NM-001-037 of the Research Division, Bureau of Medicine and Surgery. Task Order Director: Cecil W. Mann.

Following their survey of the evidence in favor of visual determinants and gravitational determinants they conclude that "gravitational (postural) factors are both logically and psychologically prior to visual factors in the determination of the perceived vertical. The framework of visual space is derived from and dependent upon the necessity of maintaining postural equilibrium against the pull of gravity."

Attempts have been made to describe the factors influencing the judgment of the visual vertical and horizontal. Jastrow (10) found that the individual could make very precise judgments of the visual vertical and visual horizontal, and believed it was due to the fact that in his common experience he is surrounded by verticals and horizontals. Koffka (11) believed that our precision in the perception of the visual vertical was due to the development of a visual spatial framework—an 'anchorage in space.' From her experiments testing this hypothesis, Neal (12) inclined to the view that postural cues were at least as important as the visual spatial framework. A recent experiment reported by Asch and Witkin (2) is considered by these experimenters as providing "striking evidence of the predominance of the visual framework over postural factors in perception of the upright."

We may except that the modification of either visual or gravitational cues, or both, will decrease the precision of vertical and horizontal perception. The control of visual cues offers no experimental difficulties. The gravitational cues may be those arising in the non-auditory labyrinth (semicircular canals and otolith bodies) or those of the non-labyrinthine proprioceptive functions set off by the stresses and strains in the muscles and tendons and the cutaneous perceptions in the trunk and limbs. Variations in these cues are possible by anaesthetisation, excision, lesion, stimulation to the point of adaptation, suggestion, or by variation in the muscle and cutaneous stresses and strains by modification of external and internal pressures. In this investigation, modification of muscle and cutaneous stresses and strains will be made by the use of padded and non-padded seats and a comparison will be made of the judgments of the postural vertical under these conditions.

Two hypotheses are introduced in this investigation:

I. In the absence of visual cues, variations of the non-labyrinthine proprioceptive and cutaneous cues will result in variation of the precision of judgment of the postural vertical.

II. The introduction of a false visual cue will modify the precision of the judgment of the postural vertical.

The investigation attempts to determine: (a) The degree of precision with which the individual, under normal conditions, can judge the visual vertical and the visual horizontal; (b) the degree of precision with which the individual, under normal conditions, can judge the postural vertical. Using these results as base lines, the investigation will then compare:

1. The precision of judgment of the postural vertical in a sitting position under modified and non-modified somesthetic conditions;

2. The precision of judgment of the visual horizontal from a position of postural tilt; and

3. The precision of judgment of the postural vertical in the presence of false visual cues.

Method

In pre-experimental trials it was found that the precision of the postural vertical reflected the method used in its determination. It was therefore decided to use three methods: (1) The method of average error; (2) the method of serial exploration, and (3) the method of single stimulus. In each method conditions of 'hard'

and 'soft' seat will be used and a comparison made of both.

Apparatus.—The apparatus consists of a lateral tilt chair, control devices to regulate the chair, and recording devices to indicate the degree of tilt.

The lateral tilt chair is made of ¾-in. pine wood and braced with iron brackets. The dimensions for the chair make possible the seating of subjects up to 200 lb. in weight. The head support is adjustable for height and for width. The *S*'s feet are maintained in position by means of strips of wood placed upon the foot rest. Push buttons controlling the direction of tilt are located on the arms of the chair. To tilt the chair to the right, *S* pushes the button on the right arm; to tilt it to the left, he pushes the button on the left arm. *S*'s arms are free to enable him to use the push buttons; his body and legs are strapped to the chair to enable him to maintain a position constant with respect to the chair as the latter is tilted.

The chair is mounted on a welded metal framework by means of a cold steel turned shaft, 37 in. above the floor. The entire mounting was designed to support 350 lb. The mounting arrangement enables the chair to be turned through 360° in a lateral plane, although in this investigation it was never tilted beyond 90° from the vertical in either direction. The chair is moved in the lateral plane by means of a reversible motor of the type used to change the pitch of propellers on large aircraft. The motor has a speed of 7000 r.p.m. and is so geared that the chair and the *S* move clockwise at 0.5 r.p.m. and counterclockwise at 0.35 r.p.m. The movement of the chair is controlled by the *S* or the *E* at the discretion of the latter.

Fitted to the chair is a removable framework of thin iron which is padded with sponge rubber. The *S* can be tested in a bare or padded chair at the discretion of the *E*. The padding is 3¾ in. in thickness on the seat and 2½ in. in thickness on the sides, back, and foot rest. The head support, in all trials, is also padded with 2½ in. of rubber on those surfaces which touch the sides of the *S*'s head and ¾ in. on the surface at the back of the head. The back of the head is supported to keep the *S*'s head tilted approximately 20° in a forward position.

The room in which the chair is located was light-proofed. It was possible to keep auditory stimuli under control although at no time was the room soundproof.

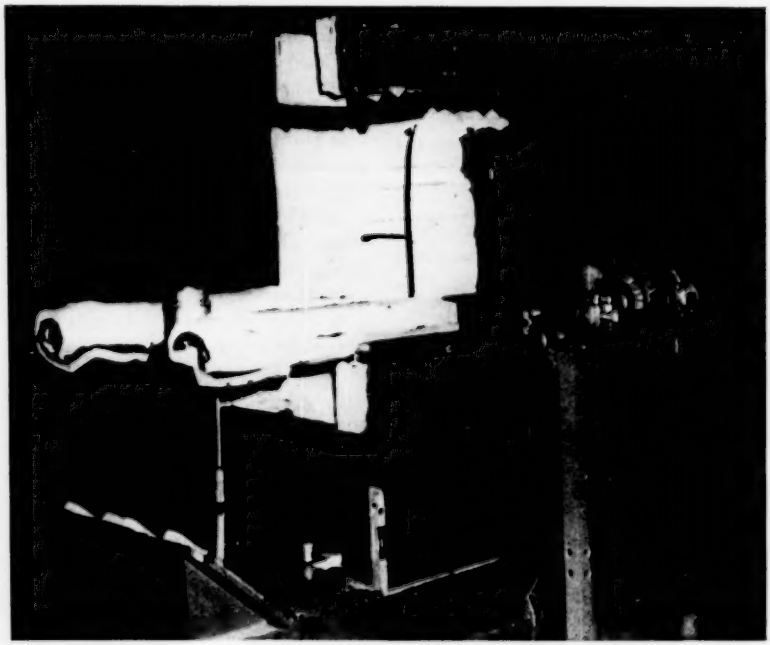

FIG. 1. Tulane lateral tilt chair

The *E* was located in a control room adjoining the chair room. The control and recording units were so constructed as to allow the *E* to (a) take control of the chair and move it either to the right or to the left; (b) give the *S* control of the chair; (c) determine the position of tilt of the chair at all times to the nearest quarter-degree, and (d) communicate with the *S* by two-way interoffice communication system.

Quantitative determination of the degree of tilt of the chair is made possible by means of a selsyn motor arrangement. The master selsyn is attached by a gear ratio of 9:1 to the shaft connecting the chair and motor. This selsyn is connected in series with a slave selsyn located in the indicating unit in the control room. To this slave selsyn is attached a needle pointer by a ratio of 9:1. Spirit levels were placed on the chair in the horizontal and vertical positions and checks were made on the selsyn pointer before and after each run.

The target consists of two 8-in. parallel strings mounted on a dull black metal frame at a distance of one in. from each other, treated with luminous paint and illuminated by an ultra-violet lamp. The target is 10 feet from *S*'s eye and 58 in. from the floor. This insures, for most *S*s, that the center of the target is at eye level. To provide constancy of illumination at all times, the target is lighted by means of a "Glo-craft" ultra-violet light unit, Model No. 106. The unit is equipped with two ultra-violet light filters, Kopp No. 51, U.V., illuminated by a 100-watt projector flood light. The light source is 34 in. from the target and strikes it at an angle of 30° above the horizontal. With pre-experimental testing, it was found that the very slight ultra-violet glow surrounding the target was of no cue value in the perception of the visual vertical. The target is moved either by *S* or *E* by means of selsyn motors. A control panel was calibrated to enable readings to be made to the nearest half degree.

Subjects.—Two groups, each of five *S*s, were used. They were college students between the ages of 20 and 27. All passed satisfactorily the 'standing test' suggested by Worchel and Dallenbach (14) as a rapid method of discriminating normal *S*s from those with vestibular dysfunction. The test consists of standing, with eyes closed, for at least 30 sec. on one foot and then on the other without overbalancing. Each *S* was given three trials on each foot.

Results

I. Judgment of the postural vertical

Three methods were used to determine the judgment of the postural:

(*a*) *Method I.*—The first method consisted of tilting the *S* *n*° laterally from the vertical and then having him return the chair to his estimate of the true vertical, first under 'hard' and then under 'soft' seat conditions.

Positions of tilt were from 5° to 90° in each quadrant in steps of 5°. One trial was given in each position, making a total of 36 trials—18 in each quadrant. The quadrants and trial positions were arranged in random order. Before each trial position, the chair was moved 15° in the opposite quadrant and then brought to the trial position. This procedure and the fact of unequal speeds in clockwise and counterclockwise movement precluded the possibility of the *S* 'counting' himself back to the true vertical. When the trial position was reached, *E* signaled the *S* and gave him control of the chair. *S* attempted to return to the true vertical and when he had made his judgment, signaled the *E* in the control booth. The *S* was allowed to swing himself to the right and to the left in making his judgment of the vertical.

The differences between the means of the judgment on hard and soft seats, as given in Table I, are not statistically significant. It will be noted that in the results for four *S*s out of five there is a statistically significant difference between the variability of the judgments of vertical on hard and soft seats. The table indicates the tendency of subjects to underestimate the vertical (minus sign denotes underestimation).

(*b*) *Method II.*—In the second method, the *S* was swept right and left across the vertical in the lateral plane and signaled the *E* when he crossed the vertical. His instructions were as follows: "This time I am going to sweep you from right to left and left to right and so on. You will

TABLE I

MEANS AND STANDARD DEVIATIONS OF INDIVIDUAL AND GROUP DEVIATIONS IN JUDGMENT OF THE POSTURAL VERTICAL OF GROUP I FOR CONDITIONS OF HARD AND SOFT SEAT, AND THE SIGNIFICANCE OF THE DIFFERENCES BETWEEN THESE TWO CONDITIONS. $N = 36$ FOR THE INDIVIDUAL CONDITIONS AND 180 FOR EACH GROUP CONDITION
(Minus sign denotes judgment in the direction of tilt.)

Subject	Mean			Variability		
	Hard	Soft	t_m	Hard	Soft	t_σ
A	0.1	0.6	0.25	1.3	2.0	4.0
B	−1.5	−2.6	0.30	2.4	4.6	3.5
C	−0.7	−0.9	0.05	1.3	2.7	3.9
D	−0.7	0.3	0.10	2.1	2.6	1.4
E	−0.9	1.3	0.10	1.7	3.1	3.3
Mean	−0.8°	−1.1°	0.60	1.9	3.4	6.5

not have control of the chair at any time. I want you to signal me with the buzzer at the exact moment you cross the vertical each time you cross it."

Each test consisted of 36 sweeps. *S* was taken from positions on the right to random positions on the left; the positions varying from 5° to 90° in both quadrants as in the first method.

The mean judgments tend to be less precise by this method than by Method I. With one exception (Subject B) the judgments show greater variability than by the use of Method I. For the group, the mean errors of judgments and the variability of errors of judgments are significantly greater (at the five percent level) for soft seat than for hard seat. All *S*s tended to anticipate the vertical.

(*c*) *Method III.*—In this method the *S* was offset *n*° from the vertical, and required to state whether he was at vertical or to the right or to the left of vertical.

The instructions were as follows: "This time you will be placed in a position either at, or very close to, the vertical. Once there, take your time and tell me over the intercom whether you are at vertical or to the right or to the left of vertical. If you are positive that your answer is correct, follow it with the word 'certain'; if you are not sure, follow it with the word 'uncertain.'"

In the trials on hard seat, *S* was offset 0.5°, 1.0°, 1.5°, 2.0°, and 3.0° in each quadrant. Since the previous results indicated a greater variability of judgment on soft seat, the *S* was offset on soft seat at 1.0°, 2.0°, 3.0°, 4.0°, 5.0°, and 6.0° in each quadrant. The procedure was to offset *S* very slowly from the vertical to the trial position. An attempt was made to move *S* by imperceptible steps to the trial position. Twelve trials were run in each position, the trial positions and the quadrants being randomized.

In Table III the percentages of correct responses for each trial position are tabulated. By "percentage of correct responses" is meant those in which the *S* was both certain and correct.

TABLE II

TABLE OF MEANS AND STANDARD DEVIATIONS OF INDIVIDUAL AND GROUP DEVIATIONS IN JUDGMENT OF THE POSTURAL VERTICAL OF GROUP I BY THE "SWEEP" METHOD FOR CONDITIONS OF "HARD" AND "SOFT" SEAT, AND THE SIGNIFICANCE OF THE DIFFERENCES BETWEEN THESE TWO CONDITIONS
(Based on 36 judgments for each subject)

Subject	Mean			Variability		
	Hard	Soft	t_m	Hard	Soft	t_σ
A	−1.9	−6.0	6.9	2.2	2.7	2.8
B	−3.3	−6.4	3.2	3.7	4.4	2.3
C	−4.0	−4.8	0.5	3.4	4.2	2.8
D	−0.8	+0.3	1.7	2.6	2.7	0.5
E	−1.8	−2.9	1.5	2.6	3.5	3.9
Mean	−2.4°	−4.0°	2.7	3.2	4.3	3.9

TABLE III

PERCENTAGE TABULATION OF CERTAIN-CORRECT RESPONSES FOR EACH ANGULAR DEVIATION IN THE METHOD OF SINGLE STIMULI FOR "HARD" AND "SOFT" SEAT FOR GROUP I
$N = 60$ for individual positions and 720 for the total.

Results on Hard Seat												
Deviation in Degrees	Left Quadrant						Right Quadrant					
	3.0	2.5	2.0	1.5	1.0	0.5	0.5	1.0	1.5	2.0	2.5	3.0
Percent Certain-Correct	100	87	63	47	23	3	13	17	47	70	90	97
Results on Soft Seat												
Deviation in Degrees	Left Quadrant						Right Quadrant					
	6	5	4	3	2	1	1	2	3	4	5	6
Percent Certain-Correct	100	73	47	23	27	17	13	20	43	37	70	100

II. Judgment of the visual vertical and horizontal

For no reason other than experimental convenience, five new *Ss* participated in the following trials. They were college students between the ages of 21 and 27, and all satisfied the clinical test for vestibular function (**14**).

The tilt chair and target were set up in a dark room. Because of the low intensity of light on the target, *S* was dark-adapted for a period of 10 min. The *S*'s head was clamped in the headrest in a vertical position throughout the experiment. The target was moved by *E* to positions of tilt varying from 6° to 15° in the right and left quadrants. Positions were chosen in random order. Each position was set twice and a total of 40 determinations was made by each *S*. The problem was to return the target to the true vertical.

Following the visual vertical determination *S* was given a three-min. rest period and then asked to make judgments of the visual horizontal. The angles of tilt used for the return to the horizontal were the same as those used in the vertical determinations, but this time were either above or below the true horizontal. The problem was to return the target to the true horizontal. The results are shown in Table IV.

From the table it will be seen that the mean errors of judgment of the visual vertical and visual horizontal are less than 1° for each *S* and that there is no significant difference between the errors of judgment of the

TABLE IV

MEAN ERRORS AND VARIABILITY IN THE JUDGMENTS OF VISUAL VERTICAL AND HORIZONTAL AND THE SIGNIFICANCES OF DIFFERENCES FOR GROUP II

(Minus values indicate underestimations of the vertical and horizontal.)

Subject	Mean Error			Variability		
	Vertical	Horizontal	t_m	Vertical	Horizontal	t_σ
A	−0.4°	−0.5°	0.30	1.5	2.9	3.89
B	0.2°	−0.3°	0.10	2.2	2.1	0.30
C	0.2°	−0.7°	0.50	2.6	2.5	0.25
D	0.2°	−0.4°	0.50	1.9	2.9	2.77
E	0.3°	0.2°	0.10	2.3	2.1	0.67
Mean	0.3°	0.4°	0.10	2.6	2.9	0.50

visual vertical and the visual horizontal. The results of two Ss indicate increased variability in the judgment of the horizontal as compared with the vertical, these differences being significant at the 0.05 level.

III. Judgment of the visual horizontal from positions of somesthetic tilt and judgment of postural vertical using false visual cues

In this part of the investigation, both target and chair were moved to the same positions of tilt. Remaining in this position, *S* was required (a) to move the target to the true horizontal, and then (b) to return himself to the true vertical. Positions of tilt from 5° to 90° in steps of 5° were used in each quadrant. One determination was made at each position in each quadrant; a total of 36 determinations. The trial positions were in random order.

The visual cue used by *S* to return himself to the postural vertical was the setting of the horizontal target made by him from the position of tilt. In the tabulation of results it was decided to use only those determinations where the error in the *S*'s horizontal setting was greater than his variable error in judging the visual horizontal from the upright sitting position. An examination of the raw data showed that two judgments of Subject A and four judgments of Subject E were less than the variable error. All other judgments of the visual horizontal made at positions of tilt exceeded the variable error of the judgments of the vertical horizontal made in the upright sitting position.

The results in Table V indicate significantly greater errors and variability when the *S*s make their judgments of visual horizontal under conditions of tilt.

In the second part of the problem, *S* was required to return himself to the postural vertical, making what use he cared to of his own setting of the visual horizontal. It is to be noted that his setting of the visual horizontal was not moved while he made his judgment of the postural vertical. The results are given in Table VI.

It will be seen that the mean errors and the variability in Group II do not differ significantly under conditions of no visual cues and false visual cues. It must be admitted, however, that the false visual cue—a pair of parallel lines eight in. long and one in. apart

TABLE V

MEAN ERRORS AND VARIABILITY IN THE JUDGMENT OF THE VISUAL HORIZONTAL FOR GROUP II WHEN MADE FROM THE TRUE VERTICAL AND FROM POSITIONS OF TILT UNDER CONDITIONS OF SOFT SEAT

Subject	Mean Error			Variability		
	Vertical	Tilt	t_m	Vertical	Tilt	t_σ
A	0.5	2.5	2.9	2.9	6.9	5.2
B	0.3	2.3	2.7	2.1	9.5	7.2
C	0.7	2.3	2.6	2.5	7.5	6.2
D	0.4	2.7	2.9	2.9	8.2	6.2
E	0.2	2.1	3.4	2.1	7.0	5.2
Mean	0.4	2.4	2.9	2.9	9.8	6.2

TABLE VI

MEAN ERRORS AND VARIABILITY OF JUDGMENT OF THE POSTURAL VERTICAL OF GROUP II UNDER CONDITIONS OF NO VISUAL CUES AND FALSE VISUAL CUES

Subject	Mean Error			Variability		
	No Visual Cues	False Cues	t_m	No Visual Cues	False Cues	t_σ
A	1.4	1.9	0.20	4.1	4.1	0.00
B	1.3	2.1	0.50	3.7	4.7	0.92
C	1.8	0.6	1.01	4.2	2.1	1.26
D	1.7	1.9	0.20	4.3	4.1	0.26
E	1.1	0.5	0.60	3.3	4.8	1.92
Mean	1.5	1.4	0.10	4.3	5.2	1.01

at a very low level of visual intensity —is not very striking. It does not resemble the 'visual framework' required by Koffka and others for their visual theory of postural judgment. Moreover, it is difficult to estimate the use made by *S* of the visual cue. At the end of a number of trials, *Ss* remarked that as they were returning themselves to the vertical, they reached a point where they realized that they had made an error in the setting of the visual horizontal. Under these circumstances it might be assumed that the false visual cue was discarded and that *S* relied entirely upon postural cues for his judgment of verticality. Until more critical conditions are set up, it is not possible to estimate the influence which false visual cues have upon the performance of this task.

Discussion

From Table I it may be seen that the mean errors of judgment of the postural vertical for hard and soft seats are not significantly different and are of the order of 1°. It would appear from these results that individuals are able to form precise judgments of verticality, and are able to return themselves to a vertical position from positions of tilt with a high degree of accuracy. This viewpoint has to be modified, however, in the light of Table II. This table gives the mean errors of judgment when the subject is swinging from right quadrant to left quadrant, and has to make a quick judgment of verticality while the chair is in motion. By this method, the differences in the mean errors of the judgment of postural vertical on hard and soft seat were statistically significant, the judgment on hard seat being the more precise.

In pre-experimental trials, it was noted that while the *Ss* were certain of their judgments of tilt at positions markedly deviating from the vertical, say at 45° or 20°, there was much uncertainty as the chair approached the vertical. This experience was commented on by all *Ss* in the experimental trials, and was checked by the time used by the *Ss* in Method I in coming to their decisions. Some *Ss* indicated that their uncertainty was occasionally so great that the expressed judgment was merely a guess.

The judgments of all *Ss* tended to conform to a normal distribution. There was a slight degree of skewness indicating the tendency to underestimate the vertical. Tilt beyond a certain angle was perceived definitely as tilt. Only as *S* approached the gravitational vertical did he experience uncertainty in his judgments. The uncertainty was most marked in an area bounded by a few degrees on each side of his perceived vertical. It is suggested that since the uncertainty of judgment increases as the vertical is approached, an 'arc of uncertainty' may be described in terms of the normal distribution of the *S*'s judgments. It is proposed to define arbitrarily the arc of uncertainty (U°) as that area including the middle 75 percent of *S*'s judgments. In each case, this will be calculated from the standard deviation of the distribution of his judgments ($U° = 1.15\sigma \times 2$).

The mean value of the arc of uncertainty (U°) for five *Ss* in Group I by each method and under conditions of hard and soft seat are shown in Table VII. It will be seen that for the group the arc of uncertainty is much greater for soft seat than for hard seat.

In Table III are shown the percentages (based upon 60 trials in each angular position) of correct responses in each angular position. It will be noted that up to 3° for hard seat and

TABLE VII

SUMMARY OF MEAN VALUES OF U° ON HARD AND SOFT SEAT FOR GROUP I USING DIFFERENT METHODS

Method	Hard	Soft
Method I	4.4°	7.8°
Method II	7.4°	9.9°
Method III		
(a) 75 percent correct responses	4.4°	10.3°
(b) Middle 75 percent of responses	3.4°	6.4°

6° for soft seat the percentage of correct responses increases as the angle of tilt increases. By interpolation on a smoothed curve it will be seen that 75 percent correct responses are given on hard seat at 2.3° L and at 2.1° R, making U° = 4.4°. By this method, U° for soft seat equals 10.3° (5.1° L + 5.2° R). Calculating the percentage based upon total responses, it was found that the middle 75 percent of responses for hard seat gave U° = 3.4°, and for soft seat, U° = 6.4°.

Under the conditions of this experiment it has been shown that while the differences in the mean judgments of the postural vertical under modified and non-modified somesthetic conditions are not significantly different, there is a trend towards greater error under modified conditions. Moreover, the variability under modified conditions is significant. The former value has been most frequently reported in the literature, but the authors of this study believe the latter to be a more useful mathematical description of the precision of judgment. For the group, it will be seen (Table I) that the standard deviation is more than twice the mean for hard seat, and more than three times the mean for soft seat. It would appear that once *S* has entered the arc of uncertainty, the position at which he stops the chair is largely a matter of chance. It is also apparent that in each method of measurement, the variability is significantly greater under soft seat than under hard seat conditions. The results, therefore, support the hypothesis that in the absence of visual cues the modification of non-labyrinthine proprioceptive cues results in a decrease of the precision of judgment of the postural vertical.

It is important to note the marked tendency for errors of judgment to fall in the direction of lateral tilt. This possibility was not considered when the experiment was designed, for it was believed to be essentially an investigation of the modification of non-labyrinthine cues. In view of this tendency, however, there is more than a possibility that adaptation to the position of tilt is a factor in the judgment of the postural vertical. It is possible that such adaptation may be derived from the kinesthetic-cutaneous indicators, or that it may involve vertibular function. Investigations of these possibilities are now in progress.

The mean errors of judgment of the visual vertical and the visual horizontal are less than 1°. Both the mean and the variability of judgment are greater when judgments are made from positions of angular tilt than when made from the postural vertical (Table V). When, however, a false horizontal target is presented it does not appear to increase the error in the judgment of the postural vertical. When visual and postural cues conflict, the postural cues appear to be the more decisive in the judgment of the postural vertical. It has been pointed out that the false visual cue in this investigation can hardly be compared with the 'visual frame of reference' adopted by Köhler. It is, however, not out of line with the instrumental horizontal upon which the

pilot depends under conditions of night flying and overcast.

Summary

1. The precision of judgment of the postural vertical from lateral tilt positions in the absence of visual cues is of the order of 0.8° with a standard deviation of 1.1°.

2. The precision of judgment is significantly decreased when the non-labyrinthine proprioceptive cues are modified by the introduction of a well padded seat.

3. There is a marked tendency for the mean error of judgment to be in the direction of lateral tilt.

4. The precision of the judgment of the visual vertical when the individual is seated in a vertical position is of the order of 0.3° with a standard deviation of 2.6°; of the visual horizontal, of the order of 0.4° with a standard deviation of 2.9°.

5. These values are significantly increased when the judgment is made in a position of lateral tilt.

6. The visual target, when placed in positions deviating not more than 15° from the horizontal, does not significantly affect the judgment of the postural vertical.

(Manuscript received October 22, 1948)

References

1. Arndts, F. Zur Frage nach der Lagewahrnehmung dienenden Sinnesfunctionen. *Pflüg. Arch. ges. Physiol.*, 1924, **204**, 539–540.
2. Asch, S. E., & Witkin, H. A. Studies in space orientation: I. Perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 325–337.
3. Birren, J. E. Static equilibrium and vestibular function. *J. exp. Psychol.*, 1945, **35**, 127–133.
4. Bourdon, B. La perception de la verticalité de la tête et du corps. *Rev. Phil.*, 1904, **57**, 462–492.
5. Dusser de Barenne, J. G. The labyrinthine and postural mechanisms. In: Murchison, C. *Handbook of general experimental psychology*. Worcester, Mass.: Clark Univ. Press, 1934, 204–246.
6. Garten, S. Über die Grundlagen unserer Orientierung im Raum. *Abhandl. Maths-Phys. Klasse Sachs. Akad.*, 1920, **36**, 431–510.
7. Gibson, J. J., & Mowrer, O. H. Determinants of the perceived vertical and horizontal. *Psychol. Rev.*, 1938, **45**, 300–323.
8. Gibson, J. J., & Radner, M. Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *J. exp. Psychol.*, 1937, **20**, 454–467.
9. Graybiel, A. Disorientation of pilots. *Contact*, 1945, **5**, 424.
10. Jastrow, J. On the judgment of angles and positions of lines. *Amer. J. Psychol.*, 1893, **5**, 214–248.
11. Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt Brace, 1935.
12. Neal, E. Visual localization of the vertical. *Amer. J. Psychol.*, 1926, **37**, 289–291.
13. Vinacke, W. E. Illusions experienced by aircraft pilots while flying. *J. aviat. Med.*, 1948, **18**, 308–325.
14. Worchel, P., & Dallenbach, K. M. The vestibular sensitivity of deaf-blind subjects. *Amer. J. Psychol.*, 1948, **61**, 94–99.

THE APPARENT LENGTH OF TILTED LINES

BY WALTER C. SHIPLEY, BARBARA M. NANN, AND MARY JANE PENFIELD

Wheaton College, Norton, Massachusetts

This experiment sought an answer to the question: as the angle of an oblique line is systematically varied from the near-horizontal to the near-vertical, is there any consistent change in its apparent length? From the facts of the horizontal-vertical illusion[1] it might be predicted that apparent length would increase as the angle approaches the vertical, perhaps even proportionately with the vertical component. The experiment, using the method of *paired comparisons*, was designed to test this prediction.

Subjects, Apparatus, and Procedure

Subjects.—The *Ss* were 48 women college students.

Apparatus.—The apparatus consisted of two black lines—3 by ¼ in. cardboard strips—centered on small, rotatable white-cardboard discs, mounted symmetrically on a larger one, as shown in Fig. 1. The smaller discs, which were 6¼ in. in diameter, were spaced ¼ in. from the edge of the larger, 12⅝ in. disc, and ¼ in. from each other. The apparatus was mounted vertically, one in. above the edge of a table 30 in. high and 30 in. wide, in such a way that none of the supporting rods and clamps were visible. *S* sat seven ft. from the apparatus, which she viewed across the table. Illumination was provided by two shaded 60-watt bulbs, one above, and the other to the right of the apparatus. The bulbs, only one of which was turned on at any given time, were located 8 in. in front of, and 12 in. radially from the center of the large disc.

Procedure.—Seven different angles of line, differing by 15-degree steps, were used. The angles can be thought of conveniently as diameters on the face of a clock, ranging by half-hour steps from 3:00 o'clock (horizontal) through 6:00 o'clock (vertical), as shown in the circular insets in Fig. 2. Each angle was compared equally often with each of the others by means of the method of paired comparisons. In each comparison *S* simply indicated the line which appeared longer, forcing a judgment when in doubt. A small screen was placed between *S* and the apparatus while the angles were being changed.

Each *S* was given four series of 21 comparisons, with the lines side-by-side in two of the series, and one-above-the-other in the remaining two. As a precaution against inequality in the actual lines, a given line was presented equally often in each of the four positions—right, left, top, and bottom—and as a precaution against shadows, the light-source was maintained equally often from above and from the side. These precautionary measures were each distributed equally for any given *S*, and rotated among *Ss*. In addition, eight serial orders of presentation—four of them the reverse of the other four—were distributed among the *Ss*.

Results

Each *S* judged the seven angles four different times in comparison with each of the remaining six, making 24 judgments for each angle and 84 judgments in all. Thus, the 48 *Ss* yielded a total of 4032 judgments, 1152 involving each angle.

The general relationship between angle of tilt and apparent length of line is given in Table I, which shows for each angle the mean number of times that the line was judged longer.

TABLE I

Mean Number of 'Longer' Judgments for Each Angle

Angle	Mean Judgments of Longer	σ
3:00 o'clock	3.58	3.94
3:30 o'clock	5.79	2.68
4:00 o'clock	11.02	3.03
4:30 o'clock	15.33	2.94
5:00 o'clock	17.06	3.18
5:30 o'clock	16.48	2.61
6:00 o'clock	14.75	3.95

[1] For a recent discussion and bibliography of this illusion see Finger and Spelt (1).

TABLE II

THE *t*-VALUES FOR DIFFERENCES BETWEEN MEAN 'LONGER' JUDGMENTS FOR PERTINENT NON-ADJACENT ANGLES

Angles Compared	Mean Difference	*t*-value
5:00 and 6:00	2.31	2.88
4:30 and 6:00	0.58	0.67
4:00 and 6:00	3.73	3.75
4:30 and 5:30	1.13	1.72

(Had all angles produced phenomenal equality the expected mean would be 12.00 for each.) These data are presented graphically in Fig. 2, which also shows *t*-values for differences between adjacent angles. Additional *t*-values for differences between pertinent non-adjacent angles are given in Table II. All *t*-values were obtained by the method of direct differences. Values greater than 2.68 are significant at the .01 level, and those greater than 2.00, at the .05 level.

From Fig. 2, considered in conjunction with Tables I and II, it will be seen (1) that all of the tilted lines are phenomenally greater, on the average, than the horizontal, (2) that apparent length increases progressively from the horizontal to the angle represented by 5:00 o'clock, and (3) that two of the angles, 5:00 and 5:30, give rise to apparent length significantly greater than that produced by the vertical. This last point is further supported by an analysis of those comparisons involving only these two angles, respectively, and the vertical. The 5:00 o'clock—6:00 o'clock comparison yielded a mean of 2.42 judgments favoring 5:00 o'clock, and only 1.58 favoring 6:00 o'clock, and the 5:30–6:00 o'clock comparison, a mean of 2.56 favoring 5:30 and only 1.44 favoring 6:00. The mean difference of 0.84 in the former comparison is significant at the .05 level, and that of 1.12 in the latter, at the .01 level.

DISCUSSION

The experiment was directed toward testing the prediction that the apparent length of a tilted line increases with its vertical component. The facts, however, showed the hypothesis to be inadequate, since certain of the tilted lines appeared longer than the vertical. In the present experiment the maximum coincided with the 5:00 o'clock angle, but since only a few angles were sampled, and certain of the obtained differences —particularly that between 5:00 and

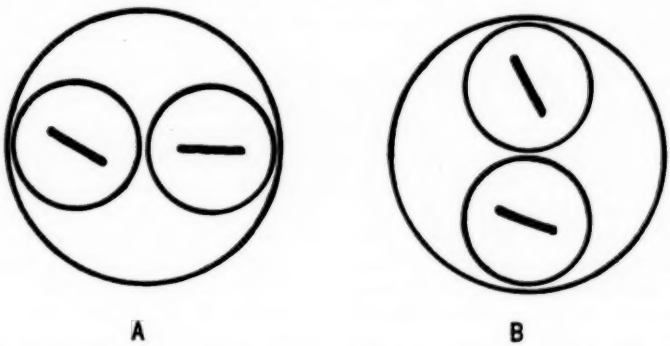

FIG. 1. Drawings of the apparatus for comparing angles of tilt. A shows the two comparison lines in the side-by-side arrangement, and B, in the top-bottom arrangement.

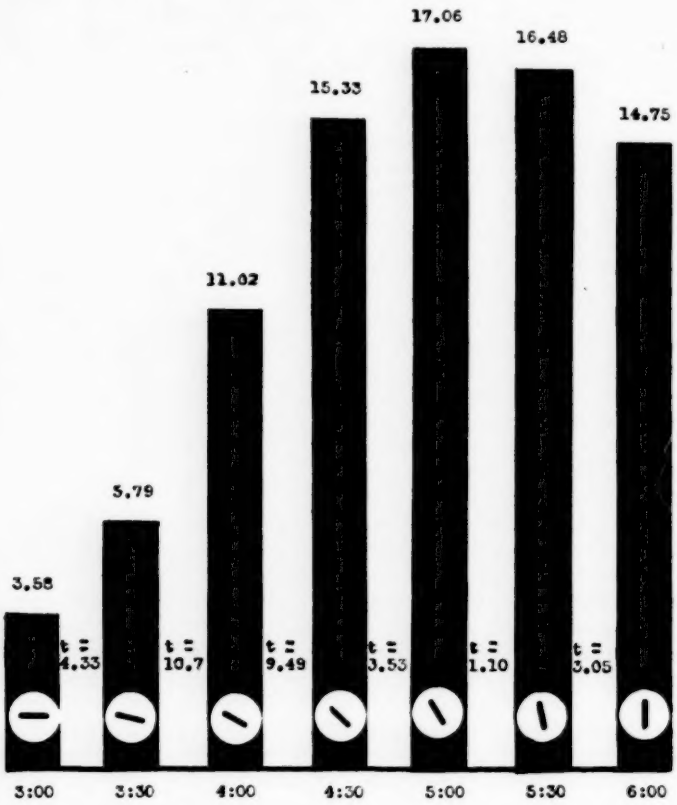

FIG. 2. Mean number of 'longer' judgments for each angle of tilt. Mean judgments are indicated above the columns, and angles of tilt, below and in the inserts. The *t*-values for mean differences between adjacent angles are shown between columns; values greater than 2.68 are significant at the .01 level, and those greater than 2.00, at the .05 level.

5:30—were not statistically significant, it can only be concluded that the 'true' maximum falls somewhere between 4:30 and 6:00. The fact that, as the angle is varied from the horizontal toward the vertical, the line progressively *increases* in apparent length up to some point in the region of 5:00 o'clock and then *decreases*, would suggest that two factors are combining to produce the effect, namely, (1) the vertical component, and (2) the deviation of the angle from the nearest major axis, i.e., either the horizontal or the vertical.

The finding that maximal phenomenal length occurs in the region of the 5:00 o'clock angle should be interpreted with the following in mind. (1) It holds only for group averages; 13 of the 48 *S*s showed a maximum at 6:00, and one, at 3:00. (2) The phenomenal length of the oblique, when taken at its maximum, exceeds that of the vertical by only a fraction of the amount that the phenomenal

length of the vertical exceeds that of the horizontal (the horizontal-vertical illusion). (3) The finding may prove to be of limited generality, being dependent upon a number of factors, such as length and relative width of line.

Summary

Forty-eight *Ss* judged seven different angles of line for apparent length, using the method of paired comparisons. The angles, regarded in terms of the face of a clock, ranged by half-hour steps from 3:00 o'clock through 6:00 o'clock. Mean judgments showed the apparent length of line to increase progressively from the horizontal (3:00 o'clock) to the angle represented by 5:00 o'clock, and then to decrease to the vertical. The occurrence of maximal phenomenal length in the region of 5:00 o'clock suggests the joint operation of two illusory factors: (1) the vertical component and (2) the deviation of the angle from the nearest major axis (i.e., either horizontal or vertical).

(Manuscript received August 24, 1948)

Reference

1. Finger, F. W., & Spelt, D. K. The illustration of the horizontal-vertical illusion. *J. exp. Psychol.*, 1947, 37, 243–250.

THE RELATIVE DIFFICULTY OF THE NUMBER, FORM, AND COLOR CONCEPTS OF A WEIGL-TYPE PROBLEM[1]

BY DAVID A. GRANT, OMER R. JONES, AND BILLIE TALLANTIS

University of Wisconsin

The purpose of this investigation was to determine the relative difficulty of the three sorting caregories, color, number, and form, of the Wisconsin Card Sorting Test (1, 3). An attempt was also made to determine some of the characteristics of the abstraction process which brought about the obtained differences in difficulty.

The Wisconsin Card Sorting Test combines features of the Weigl (17) and the Vigotsky (4, 5, 16) tests with certain principles derived from the experiments of Harlow and his coworkers on brain-damaged monkeys (14, 18). The objective quantitative scoring of different aspects of performance on this test permit considerable analysis of the *S*s' responses.

The relative difficulty of form, color, and number concepts[2] has been studied by Heidbreder (6, 7, 8, 9, 10, 11) in the 'concept-formation' situation and also in the card-sorting situation (12, 13). Her general hypothesis, although formulated in terms of 'concept attainment,' might well be expected to apply in the Weigl type of abstraction problem. From Heidbreder's hypothesis and experiments, it might be predicted that the form category of the Wisconsin Card Sorting Test would be easiest, that the number category would be most difficult, and that the color category would be intermediate in difficulty.

[1] This study was supported in part by the Graduate Research Committee of the University of Wisconsin from special funds voted by the State Legislature for 1947–48. The hospitality of the Department of Psychology of Northwestern University facilitated final preparation of the manuscript during the summer of 1948.

[2] The term 'concepts' will be used only in Heidbreder's sense (6, p. 10), namely, as denoting a logical construct without reference to the behavior of an experimental subject.

PROCEDURE

General Procedure.—The general procedure for the WCST has been described (1, 3). Briefly stated, there are 64 response cards, each containing from one to four figures of a single color. There were four colors, red, green, yellow, and blue, and four figures: stars, crosses, triangles, and circles: Thus each card could be sorted according to *color*, or *form*, or *number* of the figures. In the test situation a gray board with four double compartments was placed before the *S*. Four stimulus cards, one red triangle, two green stars, three yellow crosses, and four blue circles, were placed from left to right in the upper halves of the four double compartments. The *S* was then asked to sort the 64 response cards into the four compartments under the stimulus cards. The *S* was informed that he would be told each time whether he had been 'right' or 'wrong' in the way he sorted. These were the only instructions. Initially the category *color* was arbitrarily called correct, but after the *S* had sorted 10 consecutive cards correctly according to *color*, the *E* shifted the correct category to *form*. The *S* had no cue for the shift other than *E*'s statements of 'right' or 'wrong' as each card was sorted. After 10 correct *form* responses, the correct category was shifted to *number*, and then back to *color*, then to *form*, and finally back to *number* again.

Experimental Design.—In the standard procedure above, the correct categories are successively *color, form, number, color, form, number*. There are, however, 24 such arrangements in which one category does not appear twice in succession or before all other categories have appeared. In this experiment all 24 of these sequences were used. Four *S*s were run through each sequence, making 96 *S*s in all who completed the experiment. The performance on each category could thus be evaluated independent of ordinal effects of practice and transfer.

The *S*s were men and women students who volunteered from classes in elementary psychology at the University of Wisconsin during the summer of 1947.

Results

Each *S*'s sorting responses on each successive category were scored in terms of: (a) total errors; (b) perseverative errors; (c) non-perseverative errors; and (d) total correct responses. A *perseverative error* was scored for each card incorrectly sorted according to the category just previously counted correct. A *non-perseverative error* was any error which could not be classified as a perseverative error. The *total correct* on each successive category consisted of the number of cards sorted correctly including the last 10 'criterion' cards.

The total errors, perseverative errors, non-perseverative errors, total correct responses, and the perseverative errors on following categories are given in Table I for each of the six successive sorting categories of the experiment. Each score represents the total for 32 *S*s. It will be noted that, in general, all the error scores decrease after the second sorting, while the total correct remains more or less constant throughout the experiment, decreasing only slightly. Thus, the increased efficiency with which the *S* attacks the test consists primarily of the elimination of perseverative and non-perseverative errors and secondarily of the greater efficiency in actual acquisition of correct modes of response.

The largest number of total errors was made on the *color* category, and the least number of total errors was made on the *number* category. The analysis of variance (**2, 15**) of total error scores, however, failed to reveal a significant difference between *color*, *form*, and *number* categories. This analysis is summarized in column A of Table II.

A summary of the analysis of variance tests is given in Table II for each of the five scores considered in the experiment. Only the *F*-ratios are

TABLE I

Total Errors, Perseverative Errors, Non-perseverative Errors, Total Correct, and Perseverative Errors, on Following Category Scores for the Number, Form, and Color Concepts of the WCST

Score	Sorting Category	Successive Sorting Categories					
		I	II	III	IV	V	VI
A. Total Errors	Number	76	251	192	124	99	107
	Form	72	254	252	163	115	159
	Color	90	334	297	202	138	157
B. Perseverative Errors	Number	—	179	104	86	74	62
	Form	—	138	148	91	68	92
	Color	—	208	176	114	138	157
C. Non-perseverative Errors	Number	76	72	88	38	25	45
	Form	72	116	104	72	47	67
	Color	90	126	121	88	59	50
D. Total Correct Responses	Number	343	463	422	430	399	373
	Form	367	492	513	438	395	442
	Color	338	505	507	463	475	431
E. Perseverative Errors on Following Category	Number	—	183	207	117	96	93
	Form	—	196	133	88	82	73
	Color	—	146	89	86	43	95

TABLE II

SUMMARIES OF ANALYSES OF VARIANCE FOR TOTAL ERRORS, PERSEVERATIVE ERRORS, NON-PERSEVERATIVE ERRORS, TOTAL CORRECT RESPONSES, AND PERSEVERATIVE ERRORS, ON FOLLOWING SORTING CATEGORY

Source of Variation	*df*	*F*-ratios				
		A. Total Error Scores	B. Perseverative Errors	C. Non-Perseverative Errors	D. Total Correct Responses	E. Perseverative Errors on Following Category
(1) Sorting category	2	2.91	2.53	2.25	4.55*	3.91*
(2) Initial-final sorting	1	8.72**	3.00***	11.02**	1/1.38	53.79***
(3) Initial-final × category	2	1/6.76	1/5.99	1/13.78	1/1.46	1.28
(4) Sorting order (1-6)	4	9.26***	22.03***	1/1.40	11.28**	9.33***
(5) Sequences	23	2.32**	2.71***	1.46	4.15***	2.71***
(6) Pooled error term	111	1.15	1/1.01	1/1.06	1.22	1/1.06
(7) Individuals within sequences	72	3.06***	3.28***	1.90***	5.34***	3.28***
(8) Error (Individuals within sequences × other factors)	360	(61.13)	(18.74)	(22.03)	(25.93)	(18.74)

Numbers in parentheses are error mean squares.
* Significant at 5% level.
** Significant at 1% level.
*** Significant at 0.1% level.

presented, except that the *error* mean square appears in parentheses at the foot of each column (Row 8). These error mean squares are the denominators for the *F*'s. The breakdown of the analyses is fairly complex and merits a word of explanation. In row 1 are given the *F*'s for variation between sorting categories, the crucial test for this experiment. In row 2, the variation tested is that between the first and second occurrence of the categories; i.e., the difference between the first three successive sorting categories and the second three sorting categories, Row 2, involves a learning or transfer effect. In row 3 are evaluated the interactions between the row 1 and row 2 factors. No interaction was apparent for any score. Row 4 contains the tests for order of the six successive sorting categories minus the variation evaluated in row 2. Significant *sorting order* variation may be interpreted as general learning or transfer from one concept to another. Row 5 evaluates the variation between the 24 sequences or arrangements. (Without exception the variation between sequences could be accounted for in terms of the variation *between Ss within sequences*, as is shown by the fact that the Row 5 entries are uniformly of the same order of magnitude as the row 7 entries.) Row 6 contains the *F*'s for the variance estimate, which is a sort of pooled error term from the 24 row by 6 column table formed by the data summarized for the 24 sequences. By theory, since a complete set of permutations was used, this should give a 'pure' error estimate, and for each score this seems to be true, since none of the *F*'s in Row 6 is significant. Row 7 gives the *F*'s for consistent individual variation within the 24 sequences. High significance in each instance attests to the reliability of the scores. Row 8 gives the actual

error estimates used. They are made up of variation remaining after the variation attributable to the factors from rows 1 to 7 has been removed.

In section B of Table I it is shown how the perseverative errors decrease through the successive sorting categories. There are uniformly more perseverative errors while the *S*s try to solve the *color* category, but the analysis in column B of Table II shows that this difference is not significant.

Again on non-perseverative errors the *color* category appears to be more difficult and the *number* category is easiest in Table I, but these differences do not stand up in the analysis of variance of Table II.

The total correct responses (Section D of Table I), indicating the number of reinforcements or reminders needed to confirm or establish a consistent response, are greatest for the color category and least for the number category, again indicating the greater ease of learning to sort for number. This difference stands up in the analysis of variance (Column D, Table II), although it is significant at only the five percent level.

The perseverative errors on the following category are presented in section E of Table I. This score shows the degree to which each sorting category or concept perseverates at all stages of the experiment. (In contrast, the *perseverative error* score of Section B shows how much perseveration hampers the learning of a given response.). In Table I it is clear that the number response perseverates more than the other responses. The color response shows the least perseveration. The differences between concepts in degree of perseveration are significant at the five percent level, as is shown by the *F* of 3.91 in column E of Table II.

Discussion

The results of this experiment showed that the *S*s were able to sort selectively for *number* most easily and for *form* next most easily while sorting selectively for color was most difficult. The only scores on which the differences between sorting categories were statistically significant, however, were the total correct responses and the perseverative errors on the succeeding category. In other words, it required less reinforcement to acquire the number-sorting response and, once learned, the response of sorting for *number* tended to perseverate more than the *form* and *color* sorting responses.

The *relative* ease of acquisition of the *number* sorting response can thus be described in terms of two factors which were operating. First, it required less reinforcement or confirmation to establish the response, and secondly, perseveration of *number* sorting, being greater than that for *form* and *color* sorting, would interfere relatively more with the establishment of those responses than perseveration from *form* and *color* would interfere with *number* sorting.

The findings of this experiment seem to disagree with those of Heidbreder (**6, 7, 8, 9, 10, 11**), but Heidbreder has shown that there are obvious reasons why this disagreement may be only superficial (**12, 13**).[3] The form, color, and number concepts used are not the same as those of Heidbreder, and the *S*'s task is not the same. The ways in which the concepts and tasks varied can be specified in greater detail.

The most obvious differences between the present concepts and those of Heidbreder lie in the *number* con-

[3] Including a personal communication to the present authors.

cepts. The WCST number concepts are more restricted than Heidbreder's number concepts. Furthermore, they involve form or configurational aspects. The 'number' *one* consisted of a single figure always centered; *two* consisted of two figures, one in the upper left-hand corner and the other in the lower right-hand corner; *three* consisted of three figures forming an inverted equilateral triangle; and *four* consisted of four figures forming a square. Thus, *S*s could sort according to configuration rather than number, and verbal reports indicated that some of them did. Another fact, also, the arrangement of stimulus cards in conventional order, *one*, *two*, *three*, and *four*, from left to right across the board, would tend to make the WCST number concepts easier than those of Heidbreder. In her terms, the WCST number concepts are more perceptually accessible than hers. Research is currently being carried out to clarify the number-sorting behavior of *S*s in the WCST situation.

The differences between the task facing Heidbreder's concept-formation *S*s and the task facing the card-sorting *S*s, in her later experiments, have been pointed out (**12**, **13**). It should be added that in the WCST each sorting category is learned separately, and the *S* must spontaneously shift his mode of response after each category is learned. In Heidbreder's concept-formation experiments the different modes of response are built up simultaneously, and the shifts occur on each response. It is quite possible that the WCST accentuates difficulties due to perseveration, and if, as appears likely, the number responses are perseverated more than the form or color responses, the difference in findings may arise in part from that fact.

It is not claimed that the present findings are in conflict with Heidbreder's hypotheses. In fact, it is predicted that current research on the WCST situation will tend to confirm her hypotheses. In this respect her theory has served the Wisconsin research program as a basic tool to analyze our concepts and to clarify them. Such frames of references are greatly needed in this area. where researchers must still lift themselves by their own boot-straps.

(Manuscript received August 31, 1948)

References

1. Berg, Esta, A. A simple objective technique for measuring flexibility in thinking. *J. gen. Psychol.* (In press)
2. Grant, D. A. The latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.*, 1948, **45**, 427–442.
3. Grant, D. A., & Berg, Esta A. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *J. exp. Psychol.*, 1948 **38**, 404–411.
4. Hanfmann, Eugenia, & Kasanin, J. A. Method for the study of concept formation. *J. Psychol.*, 1936, **3**, 521–540.
5. Hanfmann, Eugenia, & Kasanin, J. Conceptual thinking in schizophrenia. *Nerv. and Ment. Dis. Monogr.*, 1942, **27**, 3–113.
6. Heidbreder, Edna. Toward a dynamic theory of cognition. *Psychol. Rev.*, 1945, **52**, 1–22.
7. Heidbreder, Edna. The attainment of concepts—a psychological interpretation. *Trans. N. Y. Acad. Sci.*, II, 1945, **7**, 171–188.
8. Heidbreder, Enda. The attainment of concepts: I. Terminology and methodology. *J. gen. Psychol.*, 1946, **35**, 173–189.
9. Heidbreder, Edna. The attainment of concepts: II. The problem. *J. gen. Psychol.*, 1946, **35**, 191–223.
10. Heidbreder, Edna. The attainment of concepts: III. The process. *J. Pyschol.*, 1947, **24**, 93–138.
11. Heidbreder, Edna. The attainment of

concepts: IV. Regularities and levels. *J. Psychol.*, 1948, **25**, 299–329.

12. HEIDBREDER, EDNA, & OVERSTREET, PHOEBE. The attainment of concepts: V. Critical features and contexts. *J. Psychol.*, 1948, **26**, 45–69.
13. HEIDBREDER, EDNA. The attainment of concepts: VI. Exploratory experiments on conceptualization at perceptual levels. *J. Psychol.*, 1948, **26**, 193–216.
14. SETTLAGE, P., ZABLE, MYRA, & HARLOW, H. F. Problem solution by monkeys following bilateral removal of the prefrontal areas: VI. Performance on tests requiring contradictory reactions to similar and to identical stimuli. *J. exp. Psychol.*, 1948, **38**, 50–65.
15. SNEDECOR, G. W. *Statistical methods.* Ames, Iowa: Collegiate Press, 1946.
16. VIGOTSKY, L. Thought in schizophrenia. *Arch. Neur. & Psychiat.*, 1934, **31**, 1063–1077.
17. WEIGLE, E. On the psychology of the so-called processes of abstraction. *J. abnorm. soc. Psychol.*, 1941, **36**, 3–33.
18. ZABLE, MYRA, & HARLOW, H. F. The performance of rhesus monkeys on series of object-quality and positional discriminations and discrimination reversals. *J. comp. Psychol.*, 1946, **39**, 13–23.

INVOLUNTARY BLINK RATE AND ILLUMINATION INTENSITY IN VISUAL WORK *

BY MILES A. TINKER

University of Minnesota

It is desirable to evaluate the techniques of measurement that are employed in studying the influence of changes in illumination intensity on ease of seeing. One of the techniques much used for this purpose is rate of involuntary blinking. The purpose of the present study is to measure blink rate while doing normal reading under a very low and under a relatively high level of illumination intensity.

Procedure

The specifications of Luckiesh, Guth, and Eastman (1) and of Luckiesh and Moss (3, pp. 102 ff.) for use of the blink technique were followed in detail. The material to be read was a section of H. G. Wells' *Outline of History*, which was used by Luckiesh and his co-workers in most of their work. This material is of uniform interest and of low and uniform emotional value. The book is printed in 10-point type with two-point leading in a 23-pica line width on a good quality rough surface near white paper.

The experiment was carried out in a light laboratory which provides well diffused illumination at various intensity levels. Forty-two university students with normal or adequately corrected vision served as readers for two experimental sessions of one hour each. The reading material (book) was placed upon an inclined stand at the normal reading distance from the subject, i.e., about 14 in. The side wall, painted with a flat cream tint with a reflection factor of about 55 percent, served as a background. This wall was about 10 feet from the subject.

The direct method of observing and counting eye blinks by a skilled experimenter was employed. On arrival at the laboratory the subject was adapted for five min. to the illumination used. During this time he was told that he would be asked to do some ordinary reading, and that we wanted him to read the material just as he was accustomed to read material of that kind, that is, to read it just as he would if he were reading at home as a leisure activity. Then the subject was told to read the material at his usual rate and with his usual attitudes, to start at the indicated place, and to continue until stopped.

The odd-numbered subjects read under two footcandles of illumination at the first session and under 100 footcandles at the second session. The even-numbered subjects read under 100 footcandles at the first and under 2 footcandles at the second session. (Luckeish and Moss (2) employed 1, 10, and 100 footcandles. The 2 and 100 footcandles used here approximate their lower and upper limits.) If the subjects asked what the experiment was about, they were merely told that we wanted to know how well they liked to read under different kinds of light. At the end of the second session, as a gesture, each subject was asked which light he preferred, the dim or the bright. (Twelve preferred the 2, and 30 the 100, footcandles.)

Each subject read for 55 min. under the 2 footcandles and for the same time under the 100 footcandles. This is five min. less than the Luckiesh and Moss time (2) of 60 min. The illumination was periodically checked with a recently calibrated General Electric footcandle meter. Blinks were tabulated during the first five and last five min. of reading at each session. The experimenter sat off to the side and slightly back from the readers' visual field. The only activity of the subject outside of reading was to turn the pages of the book. Every effort was taken to maintain normal reading attitudes.

Results and Discussion

The basic data of the experiment are listed in Table I. The means and standard deviations for the initial and final five min. of reading, the correlation between blinks for initial and final five min., and the differences between blinks in initial and final five min. with critical ratios are given. The correlation between initial and final number of blinks may be considered a measure of reliability. It is

* The writer is grateful to the University of Minnesota Graduate School for a research grant to finance this study.

TABLE I

INFLUENCE OF ILLUMINATION INTENSITY UPON RATE OF BLINKING DURING 55 MIN. OF READING

Footcandles	First 5 Min.		Last 5 Min.		r	Diff.	% Diff.	$D/\sigma_{Diff.}$
	Mean	SD	Mean	SD				
2	25.76	23.15	33.90	27.86	.87	8.14	31.6	5.60
100	26.83	19.38	35.43	24.48	.57	8.60	32.0	3.97

high (+.87) for the reading under two footcandles, but rather low (+.57) for the 100 footcandles. The latter is high enough, however, for group comparisons.

Examination of the means reveals that there is an appreciable increase in number of blinks from initial to final five min. of reading under both levels of illumination. The changes in blink rate from the beginning to the end of 55 min. of reading, in terms of the critical ratio, are both significant at the one percent level. The *t*-test of significance was also computed. The results here also indicated significant differences at the one percent level: $t = 3.76$ and $t = 2.65$ for 2 and 100 footcandles, respectively.

For two footcandles, the increase in blink rate from initial to final five min. of reading was 8.14, or 31.6 percent. Similarily for 100 footcandles, the increase was 8.60, or 32.0 percent. Thus, we find an almost identical amount of increase in blinks for reading approximately an hour under 2 and under 100 footcandles.

It may be of some interest to note the performance of those who preferred the 2 footcandles as compared to those who preferred the 100 footcandles for reading. For the 12 subjects who preferred the 2 footcandles, the mean number of blinks under 2 footcandles was 27.4 for the initial five min. and 32.8 for the final five min. with a difference of 5.4. Under the 100 footcandles the respective figures are 32.6 and 32.1, with a difference of −0.5. For the 30 subjects who preferred 100 footcandles the mean number of blinks under 2 footcandles was 25.1 for the initial five min. of reading and 34.3 for the final five min., with a difference of 9.2. While reading under 100 footcandles, the respective figures for these subjects are 24.5 and 36.7, with a difference of 12.2. Thus, for those who preferred to read under 2 footcandles, there is a suggestion that reading for 55 min. under 2 footcandles produces a greater increase in blinks from initial to final five min. than when reading under 100 footcandles. For those who preferred the 100 footcandles, however, the trend is reversed, i.e., there is a tendency to be a greater increase in blinks for reading under 100 footcandles than under 2 footcandles. When the two subgroups of 12 and 30 subjects were separated from the total group, the arrangement for having subjects alternate in starting (first observation session) with 2 or with 100 footcandles was upset. It seem reasonable to infer that more valid results are derived from consideration of the total group of 42 subjects.

It is obvious, in terms of the literature on the subject, that reading under 2 footcandles is a more severe visual task than reading under 100 footcandles. But these results reveal no

such trend. It would seem, therefore, that under these experimental conditions, rate of involuntary blinking is not a valid criterion of ease of seeing in reading.

Luckiesh and Moss (2) obtained the number of blinks during the first and last five min. of an hour of reading under 1, 10, and 100 footcandles. Eleven subjects were used. Their 1 and 100 footcandles constituted a lighting situation comparable to the 2 and 100 footcandles in this study. In their study, Luckiesh and Moss obtained mean scores of 35 and 60 blinks for the first and last five min. of reading under the one footcandle, and 36 and 39 blinks for the first and last five min. of reading under the 100 footcandles. This means an increase in blinks of 71.5 percent when reading 60 min. under one footcandle and only 8.3 percent increase for the same amount of reading under 100 footcandles. The trends in the Luckiesh-Moss results are not in agreement with the findings in the present study, where the increase in percent of blinks from first to last five min. of reading for 55 min. was approximately the same for the 2 and for the 100 footcandles, i.e., 31.6 and 32.0 percent.

The fact that I employed 2 and 100 footcandles while Luckiesh and Moss used 1 and 100 footcandles of illumination does not prevent a comparison of the trends in the two studies. This statement seems valid because Luckiesh and Moss found a 31.4 percent increase in blinks from first to last five min. for reading 60 min. under 10 footcandles. This was much more than they found for the 100 footcandles. I employed 42 subjects while Luckiesh and Moss used only 11. The main effect of using more subjects would be to make the mean scores more reliable. But the similarity of number of blinks during the initial five min. of reading under 1, 10, and 100 footcandles by the Luckiesh and Moss subjects (35, 35, and 36 blinks, respectively) would suggest that reliability was probably not involved.

Reasons for lack of agreement between trends in the Luckiesh and Moss (2) data and those reported here are not clear at this time. It would seem that the differences cannot be assigned to differences in experimental method and probably not to differences in size of groups studied. Apparently the source of the lack of agreement must be sought elsewhere.

Summary

1. The purpose of this study is to measure blink rate while doing normal reading under a very low and under a relatively high level of illumination intensity.

2. Forty-two subjects read for 55 min. under 2 and under 100 footcandles of light. Blinks were recorded for the first and for the last five min. of reading under each kind of illumination.

3. The number of blinks per five min. of reading increased by an almost identical amount from the first to the last five min. under the two intensities of light.

4. Under the conditions of this experiment, rate of involuntary blinking does not reflect differences in ease of seeing.

(Manuscript received November 22, 1948)

References

1. Luckiesh, M., Guth, S. K., & Eastman, A. A. The blink-rate and ease of seeing. *Illum. Engng.*, 1947, 42, 584–588.
2. Luckiesh, M., & Moss, F. K. Frequency of blinking as a clinical criterion of ease of seeing. *Amer. J. Ophthal.*, 1939, 22, 616–621.
3. Luckiesh, M., & Moss, F. K. *Reading as a visual task.* New York: D. Van Nostrand Co., 1942, pp. 428.

THE EFFECT OF A FIXATED FIGURE ON AUTOKINETIC MOVEMENT

BY RICHARD S. CRUTCHFIELD

Swarthmore College

AND

WARD EDWARDS

Harvard University

No completely satisfactory theory has yet been offered to account for the fact that a small, objectively stationary dot of light in a dark room is usually seen as moving. Theories about this phenomenon, called autokinetic movement, have been of two different kinds. The first attempts to explain the phenomenon in terms of peripheral processes, for example, eye movements. The second postulates some sort of central process.[1]

In 1879, Hoppe (8) proposed that autokinetic movement, like other illusions of visual movement, depends on the functioning of the extrinsic eye muscles. The actual causal factor was assumed to be either the innervation of these muscles or the shifts of the image of the dot on the retina as a result of the ensuing movements. In 1886, Charpentier (4) defended the second of these possibilities in spite of his experimental findings, subsequently confirmed by Marx (11) and others, which showed that large scale eye movements could not be involved. In 1928, Guilford and Dallenbach (7) studied eye movements of much smaller size by means of a photographic record and found no relationship between these and autokinetic movement.

[1] An excellent summary of the literature on autokinetic movement up to 1928 is found in Guilford and Dallenbach (7). Summaries of more recent literature may be found in Hovland (9) and Skolnick (15).

A more recent paper by Skolnick (15) contends that the Guilford and Dallenbach results are invalid because their photographic technique was insensitive and because it recorded only horizontal eye movements while the most common direction of autokinetic movement is vertical. On the basis of his own experiments Skolnick concludes that the arguments levelled against the eye movement hypothesis are invalid, and that there is considerable evidence in favor of the hypothesis.

In 1910, Carr (3) proposed strains of the extrinsic eye muscles as being one important causal factor in autokinetic movement, on the basis of observations, later confirmed by Guilford (6), that movement occurs in the direction of eye strain. Guilford, however, cast doubt on this theory because he found no correlation between strength or weakness of extrinsic eye muscles and the predominant direction of autokinetic movement for any of a series of subjects, and because he felt it more likely that the experienced eye strains were a "result of the visual phenomenon (i.e., autokinetic movement), perhaps set up in the attempt of *O* to inhibit the tendency to follow the moving light with his eyes" (6, p. 409).

Guilford and Dallenbach (7) and Guilford (6) proposed a third explanation of autokinetic movement. They considered it a case of induced

movement, resulting from the passage of the 'streaming phenomenon' across the perceived autokinetic dot. They offered phenomenological evidence for this hypothesis, and pointed out that the streams which are involved must be, unlike streams described by previous authors, such as to involve the whole retina at once.

Any theory of autokinetic movement dependent on such peripheral factors as those proposed above is made less plausible by the fact that the movement is subject to both voluntary and social control. This is shown by a series of studies (1, 2, 3, 4) culminating in the work of Sherif (12, 13, 14). In 1935, Sherif (12) found that the range and norm of an individual's judgments as to the amount of distance covered by the moving dot of light could be changed by a knowledge of the judgments of other individuals perceiving the phenomenon in the same social situation. Quite large and relatively permanent changes in the range and norm of the individual's judgments could be thus produced. Changes of up to 180 degrees in the *direction* of movement seen could also be produced in appropriate social situations.

It seems unlikely that such voluntary and social influences on autokinetic movement are mediated by peripheral processes. Moreover, Voth (16) reports success in using autokinetic movement as a method of personality measurement. Taken together these considerations argue for a central determination of the autokinetic effect.[2] The present paper offers further support for a theory of central determination and throws some light on its possible nature.

The experiment to be reported here studied the effect of the fixation of a particular visual pattern on subsequent autokinetic movement. In essence the experiment consisted of three steps. First, the prevailing direction and extent of autokinetic movement were determined for the subject. Second, a visual figure consisting of a luminous outline semicircle was fixated for several minutes by the subject, the semicircular figure being exposed either (a) on the same side as the prevailing direction of autokinetic movement, or (b) on the opposite side. Third, as soon as the fixated figure was removed the direction and extent of autokinetic movement were again determined for the subject. Comparison was then made of the character of the autokinetic movement before and after fixation of the figure.

Apparatus and Subjects

The stimulus light for autokinetic movement was made by enclosing a standard flashlight bulb in a light-tight cardboard container with a circular hole one mm. in diameter. The hole was covered with one thickness of onionskin paper to diffuse the light. An electrical timer controlled the length of exposure of the light, the exposure time being set at 2.14 sec. per trial.

It was desirable to insure ample autokinetic movement in all subjects, and it had been ascertained in preliminary work that the humming sound of a motor tended to enhance the extent of autokinetic effect and to induce it more readily in subjects who were slow to perceive the movement. For this reason, an electric motor which served solely as a sound effect was connected in parallel so that it ran during exposure of the stimulus light.

The fixation figure consisted of a luminous semicircular line (with radius of four in.) and a small spot of light located on the bisecting radius, one in. from the center of the circle. The line was five mm. in width and the spot of light five mm. in diameter. The fixation figure was cut out of a larger circular cardboard disk, which was placed in front of a light-tight box and so arranged that the disk could be rotated around

[2] Graybiel and Clark (5) also have implied a central theory when they suggest that the real problem is not why the autokinetic dot moves, but rather why more complicated perceptual situations do not yield experiences of autokinetic movement.

its center, thus enabling the fixation figure to be presented in any desired orientation. The fixation figure was covered with one thickness of onionskin paper to diffuse the light. The figure was approximately twice as bright as the autokinetic stimulus dot. The length of exposure of the fixation figure was controlled by a timer and set at 140 sec.

The subjects for the experiment were Swarthmore College students of both sexes. Since none of them had ever heard of autokinetic movement and none reported doubt of the objectivity of the perceived movement, it can be safely assumed that they were completely naive for the purposes of this experiment.[3]

Procedure

The subject was seated in the dark room 15 feet from the apparatus. He was told that he would be shown a moving dot of light and was instructed to report orally the estimated amount of the movement (in in.) and its direction. He was instructed also to draw a line on a pad of paper in front of him to represent as closely as possible the path of the perceived movement.

The subject was then given 10 successive trials with the autokinetic dot visible for 2.14 sec. per trial. An interval of roughly 20 sec. occurred between trials, during which the amount and direction of perceived movement in the previous trial were reported and recorded.

Preliminary experimentation had shown, as earlier reported by Sherif (12, 13), that a subject quickly establishes a given amount and direction of perceived movement which tends to persist from trial to trial. With virtually all our subjects in this experiment a fairly stable direction and amount of movement was established by the end of the ten trials. For all but two of our subjects the movement was either horizontal or above the horizontal.

The 26 subjects in the experiment were divided randomly into four groups with equal numbers of males and females in each group, Immediately following the 10 autokinetic trials, each subject in Group I was presented with the semicircular fixation figure so arranged that it was in the prevailing direction of his perceived autokinetic movement. That is, if he had perceived movement regularly to the right, the semicircle was placed at the right, if he had perceived movement upward, the semicircle was placed above, etc. (See Fig. 1a.) The subject was instructed to fixate the spot near the center of the semicircle, and the figure was exposed for 140 sec. Five sec. after the fixation figure was removed, the autokinetic light was again presented and a record of the movement obtained. Thereafter at 15-sec. intervals four more autokinetic trials were given. All five of these autokinetic trials were given in an identical way with the original series of ten trials before presentation of the fixation figure.

Each subject in Group II was treated in precisely the same way as those in Group I, except that the fixation figure was presented on the side *opposite* to the prevailing direction of the subject's perceived movement. (See Fig. 1b.)

In Group III, each subject used only his right eye during all autokinetic trials and only his left eye during inspection of the fixation figure. The fixation figure was placed as it was for Group I, namely, in the prevailing direction of the autokinetic movement.

The subjects of Group IV, the control group, were treated like the others in the autokinetic trials, but instead of a period of exposure of the fixation figure, they were simply required to sit in the dark room for an equivalent period.

Results

The extent of autokinetic movement in the final five trials before exposure of the fixation figure and in the first five trials thereafter is shown for all four groups in Table I.[4] The effects in the various groups can be better compared if the extent of autokinetic movement in each trial is expressed as a percentage of the amount of

[3] In view of Sherif's work, and of the notorious responsiveness of autokinetic movement to suggestion, the importance of naivete in subjects for such experiments is to be emphasized. For instance, in our own preliminary work we discovered that subjects who were acquainted with the subjective character of the phenomenon tended to give results different from those obtained from naive subjects. This suggests a possible criticism of many of the earlier studies of autokinetic movement cited above. In studies other than Voth's (16) and Sherif's (12, 13, 14), the subjects were usually either the experimenters themselves, or psychologists and other individuals familiar with autokinetic movement, whose theoretical expectations might well influence the results of their observations. (This criticism applies to some but not all of Skolnick's experiments.)

[4] Only the final five trials before exposure of the fixation figure are reported here because in some cases a stable amount and direction of movement was not yet established in the earlier trials.

TABLE I

MEAN AMOUNTS OF MOVEMENT IN INCHES FOR PRE-FIXATION AND POST-FIXATION TRIALS

Trial	Group I (N =8)	Group II (N =6)	Group III (N =6)	Group IV (N =6)
1	6.50	1.79	4.83	3.67
2	6.06	2.17	5.92	4.08
3	5.88	2.42	4.83	3.83
4	5.63	2.75	5.08	3.33
5	5.88	3.00	5.25	3.67
Fixation period				
6	1.88	0.75	2.92	3.58
7	4.38	1.21	4.25	3.54
8	5.38	1.29	5.00	3.67
9	5.38	1.58	5.50	3.67
10	6.00	1.89	5.33	4.17

TABLE II

MEAN AMOUNTS OF MOVEMENT EXPRESSED AS A PERCENTAGE OF THE AMOUNT OF MOVEMENT IN THE LAST TRIAL BEFORE FIXATION

Trial	Group I (N =8)	Group II (N =6)	Group III (N =6)	Group IV (N =6)
1	111%	60%	92%	100%
2	103	67	113	111
3	100	81	92	104
4	96	92	97	91
5	100	100	100	100
Fixation period				
6	32	25	56	98
7	75	40	81	96
8	92	43	95	100
9	92	53	105	100
10	102	65	102	113

movement in the last trial before exposure of the fixation figure. These adjusted data are shown in Table II and Fig. 2. The effect is to balance out the variation among the groups in mean absolute amount of autokinetic movement (a variation which must be presumed to be simply a function of the small size of the groups, since subjects were assigned to the four groups in a random manner), and this permits a direct comparison of the proportion of reduction in autokinetic movement in the several groups.

Groups I, II, and III show a marked drop in extent of autokinetic movement following exposure of the fixation figure. Group IV, the control group, exhibits no change in extent of movement after its period of darkness interpolated between the first set and the last set of autokinetic trials.

The reduction in Groups I, II, and III is sharpest for the first trial following the fixation figure, and on subsequent trials there is a gradual increase in extent of movement until by the fifth post-fixation trial it has returned to about the level of the trials before fixation. (Group II shows a somewhat slower recovery.) The drop in the first post-fixation trial was exhibited by virtually all sub-

FIG. 1. The semicircular fixation figure exposed (a) on the same side as the prevailing direction of autokinetic movement (indicated by the dotted line), and (b) on the opposite side

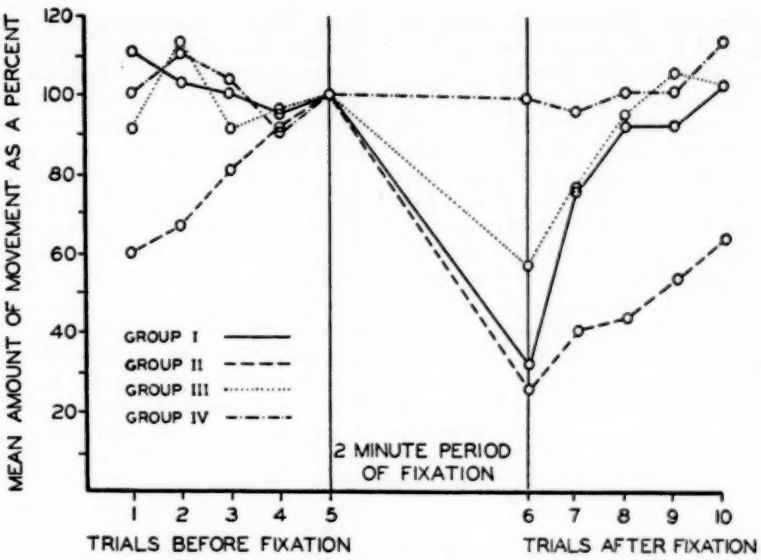

FIG. 2. Mean amounts of autokinetic movement for pre-fixation and post-fixation trials expressed as a percentage of amount of movement in the last trial before fixation

jects. Only one subject (Group II) reported a larger extent of movement in the first post-fixation trial, and this increase was only $\frac{1}{8}$ in. Two subjects (Groups I and III) reported no change in amount of movement. All the rest of the subjects reported smaller amounts of movement in the first post-fixation trial, with decreases ranging from one-half to 10 in. The individual records are shown in Table III.

TABLE III

MOVEMENT IN LAST TRIAL BEFORE FIXATION (TRIAL 5) AND FIRST TRIAL AFTER FIXATION (TRIAL 6) IN INCHES FOR EACH SUBJECT

Subject	Group I		Group II		Group III		Group IV	
	Trial 5	Trial 6	Trial 5	Trial 6	Trial 5	Trial 6	Trial 5	Trial 6
1	6	1	2	0.25	0	0	4	4
2	4	2	0	0.12	1	0.5	5	4
3	4	0	3.5	0.12	8	6	4	5
4	2	2	8	3	12	6	6	5.5
5	5	2	2.5	0	4.5	3	1	1
6	4	2	2	1	6	2	2	2
7	10	0						
8	12	6						
Means	5.88	1.88	3.00	0.75	5.25	2.92	3.67	3.58
t of diffs. between trials 5 and 6	3.68		3.04		2.52		—	

The mean differences between the last trial before fixation and the first trial after fixation yield a t of 3.68 for Group I, a t of 3.04 for Group II, and a t of 2.52 for Group III. These differences are significant at better than the one percent level, better than the five percent level, and just at the five percent level, respectively. In Group IV, the mean difference between the last trial before interpolation of the period of darkness and the first trial thereafter is only 0.09 in., and obviously not significant.

No significant variation in prevailing *direction* of movement was found in the subjects as a result of exposure of the fixation figure.

Discussion

On the basis of the above results it seems perfectly clear that a period of exposure of the fixation figure causes an immediate and temporary reduction in the amount of perceived autokinetic movement. Moreover, the fact that this reduction is found also in Group III, in which the autokinetic movement was observed with one eye and the fixation figure with the other, provides evidence for the conclusion that this reduction effect is not peripheral, i.e., retinal, but is *central* in locus. That the extent of autokinetic movement is so readily susceptible to influences which are mediated centrally argues strongly, we believe, for a theory of the central determination of the autokinetic phenomenon itself.

We can only conjecture about the possible nature of the central nervous system processes which govern the observed reduction in autokinetic movement. It is provocative to consider the possibility of explaining the results in terms of the theory of satiation offered by Köhler and Wallach (10). These experimenters propose the view that all processes in the visual cortex alter the medium in which they take place in such a fashion as to affect the subsequent occurrence of a process in the same or a nearby visual area. This alteration they call satiation. From the point of view of this theory of satiation the effect of exposing our semicircular fixation figure would be to alter the medium around the process in the visual cortex corresponding to it (the S process) in such a fashion that when a process corresponding to the autokinetic dot (the A process) occurs, movement of that A process might be expected to be hindered by the increased resistance of the satiated medium in the area surrounding the S process.

This application of the satiation theory can account very well for the reduction in extent of autokinetic movement found in Groups I, II, and III after exposure of the satiating fixation figure in the vicinity of the location of the autokinetic dot. It accounts also for the gradual recovery in amount of the autokinetic movement to its pre-satiation level, since the satiation is assumed to wear off gradually.

But the theory would also seem to call for a *difference* in the amount of the reduction effect in Groups I and II, though no such difference was actually found. For Group I the fixation figure was exposed on the same side as the prevailing direction of perceived movement, while for Group II it was exposed on the opposite side. One might expect, therefore, that the amount of reduction in Group I would be greater than that in Group II, since for the former the A process might be thought of as moving *into* the region of maximal satiation at the place of the S process, while for the latter the A process might be

thought of as moving *away* from the region of maximal satiation.

The fact that this expected difference in reduction between Groups I and II was not found can be harmonized with the satiation theory by making the assumption that the relatively long two-min. period of exposure of the fixation figure resulted in fairly intense and uniform satiation over an area sufficiently extensive for the movement of the A process to be equally impeded in any direction. This assumption can be tested experimentally by decreasing the time of exposure of the fixation figure, which would serve to reduce the general level of satiation and hence permit the establishing of a differential amount of satiation for the A process in the directions toward and away from the place of the S process. The authors will later report experiments of this nature in which the exposure time for the fixation figure is decreased to one min.

The authors believe that the results cannot be accounted for by practice in fixation. There appears to be no evidence that degree of fixation affects autokinetic movement. Also, the reduction in autokinetic movement would scarcely be expected to disappear so rapidly after the first or second post-fixation autokinetic trial, since it would not seem likely that the improved ability to fixate should disappear so quickly.

Summary

The object of the experiment reported here was to determine the effect of the prolonged fixation of a particular visual pattern on subsequent autokinetic movement in the region of the pattern. The experiment consisted of three steps. First, the prevailing direction and extent of perceived autokinetic movement were determined for each of the 26 subjects. Second, a visual figure consisting of a luminous outline semicircle was fixated for two min. by the subject, the semicircular figure being exposed either (a) on the same side as the prevailing direction of autokinetic movement for that subject, or (b) on the opposite side. Third, as soon as the fixated figure was removed the direction and extent of autokinetic movement were again determined for the subject. For some subjects the autokinetic movement was observed with one eye and the fixation figure with the other eye. A control group was given a two minute period of darkness instead of the fixation figure.

A comparison of the autokinetic movement before and after exposure of the fixation figure revealed the following:

1. The exposure of the fixation figure caused a marked reduction in extent of autokinetic movement, the reduction being immediate and temporary only.

2. This reduction was also found for subjects who viewed the autokinetic movement with one eye and the fixation figure with the other, thus implying a central rather than a peripheral mediation of the reduction effect.

3. The same amount of reduction occurred whether the fixation figure was placed on the same side as the prevailing direction of perceived movement or on the opposite side.

4. No effect on direction of autokinetic movement was found.

These results are discussed in light of the Köhler-Wallach theory of satiation in figural after-effects, and the need for certain further experiments to test this interpretation is indicated.

(Manuscript received for immediate publication April 6, 1949)

References

1. Adams, H. F. Autokinetic sensations. *Psychol. Monogr.*, 1912, **14**, 1–45.
2. Bourdon, B. *La perception visuelle de l'espace.* Paris: Schleicher Frères, 1902.
3. Carr, H. A. The autokinetic sensation. *Psychol. Rev.*, 1910, **17**, 42–75.
4. Charpentier, A. Sur une illusion visuelle. *C. R. Acad. Sci., Paris*, 1886, **102**, 1155–1157.
5. Graybiel, A., & Clark, B. The autokinetic illusion and its significance in night flying. *J. aviat. Med.*, 1945, **16**, 111–151.
6. Guilford, J. P. Autokinesis and the streaming phenomenon. *Amer. J. Psychol.*, 1928, **40**, 401–417.
7. Guilford, J. P., & Dallenbach, K. M. A study of the autokinetic sensation. *Amer. J. Psychol.*, 1928, **40**, 83–91.
8. Hoppe, J. I. *Die Scheinbewegung.* Würzburg: A. Stuber, 1879. (Not seen, cited in 7).
9. Hovland, C. I. Apparent movement. *Psychol. Bull.*, 1935, **32**, 755–778.
10. Köhler, W., & Wallach, H. Figural after-effects, an investigation of visual processes. *Proc. Amer. phil. Soc.*, 1944, **88**, 269–357.
11. Marx, E. Untersuchungen über Fixation unter verschiedenen Bedingungen. *Z. Sinnesphysiol.*, 1913, **47**, 79–96.
12. Sherif, M. A study of some social factors in perception. *Arch. Psychol.*, 1935, No. 187.
13. Sherif, M. *The psychology of social norms.* New York: Harper's, 1936.
14. Sherif, M. An experimental approach to the study of attitudes. *Sociometry*, 1937, **1**, 90–98.
15. Skolnick, A. The role of eye movements in the autokinetic phenomenon. *J. exp. Psychol.*, 1940, **26**, 373–393.
16. Voth, A. C. Individual differences in the autokinetic phenomenon. *J. exp. Psychol.*, 1941, **29**, 306–322.

THE DISCRIMINATION OF SMALL DIFFERENCES IN THE TIME OF MECHANICAL STIMULATION*

BY F. L. SMITH, A. L. SWEET,[1] AND N. R. BARTLETT[2]

The Johns Hopkins University

The present experiment was performed to get some idea of the ability of *O*s to discriminate tactually differences in the phase of vibrations. We had both basic and practical objectives in mind in doing such an exploratory experiment.

On the basic question how well *O*s discriminate phase differences in vibration, the existing literature has little to say. There have been a great many studies of vibration sensitivity. In these studies, however, the interest has been in high frequencies and in the frequencies and amplitudes of vibration which are just detectable by *O*s (**1**, **2**, **3**, **4**, **5**, **9**, **10**). None has attacked the problem of discriminating differences in phase of two vibrating objects which have the same frequency and same amplitude of vibration. Since phase is one property of vibrational stimulus and since phase of auditory stimuli has been a factor commanding considerable attention, this factor in vibrational stimuli seemed worth looking into.

As a practical matter, too, there is the question whether phase-differences in vibration might ever prove useful as a means of presenting information to, say, a machine operator. The use of tactual cues would seem to offer advantages with apparatus whose operation requires response to several simultaneous indications. If the operator's hands, in addition to manipulating the controls, can be 'built into' an instrument system, then there would be less burden upon his visual and auditory senses. These senses would, as a consequence, be more free to handle information from other channels.

So this study concerns the use of phase as a sensory cue in discriminating vibrations of low frequency. Put another way and more specifically, the study gives us measurements of sensitivity to differences in the time of mechanical stimulation of the finger tips. Measurements have been made for five different rates of stimulation.

Apparatus

A diagrammatic sketch of the apparatus used is included in Fig. 1. Two wooden ¼-in. dowels were mounted on a common horizontal axis (a) about one-third the distance from the ends facing *O*. Vertical movement was imparted to the dowels by two eccentrics (E_1, E_2) mounted on a common axle so that their speeds were identical. Changes in rate of movement were made possible by different combinations of belt-drives connecting the axle of the eccentrics with a motor. Phase relations of the dowels were changed by means of the graduations on the eccentrics and two reference pointers (P_1 and P_2). Setting eccentric E_1 to the zero position as indicated by the reference pointer P_1 permits the adjustment of eccentric E_2 simply by setting the degree phase difference desired against the pointer P_2.

The dowels were so mounted on the eccentrics that the ends used by *O* had an excursion of approximately ¾ in.

In this experiment, five different speeds were used: 50, 72, 288, 710, and 830 rpm. Speed was

* This research was supported in part by contract N5ori-166, Task Order I, between the Special Devices Center, Office of Naval Research, and The Johns Hopkins University. This is report No. 166-I-87, Project Designation No. NR-784-001, under that contract.

[1] Now at the Psychology Section, Radio Division III, Naval Research Laboratory, Anacostia, D. C.

[2] Now at the Department of Psychology, Hobart and William Smith Colleges, Geneva, New York.

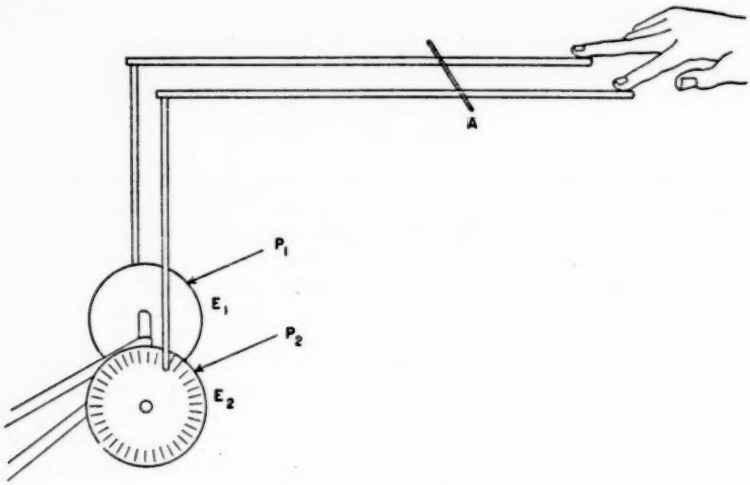

FIG. 1. Diagrammatic sketch of apparatus

regularly checked with a tachometer on the eccentric axle.

PROCEDURE

The ends of the dowels presented to the *O*s moved through a total excursion of about ¾ in. The *O* placed the first and middle figures on the ends of the dowels and reported whether the dowels were in phase or out of phase. Reaction time was measured from the time of contact of *O*'s fingers with the dowels to the beginning of his verbal response. Both reaction time and the verbal response were recorded.

As we said before, five speeds were used. At each speed, the observer had to make judgments of phase differences varying between zero and 180 degrees. Each *O* made five judgments of each phase difference for each speed. The number of settings of phase differences varied for different speeds in order to provide the greatest number of judgments in the vicinity of what, in preliminary tests, appeared to be the threshold.

Five male subjects were used. Since the right hand was the preferred hand for all five *O*s, it was the hand used in all measurements.

There were two measures of *O*s' judgments. One was reaction time, which always was recorded whether the judgment was correct or not. The other was the percentage of judgments that the dowels were out of phase. To display the results graphically, all phase differences have been converted from degrees (in which the apparatus was calibrated) to time-differences.

RESULTS

Results are presented in Table I. The table gives the data for each of the five speeds used. Each value in the table is for 25 readings (five readings for each of five *O*s). An inspection of Table I shows the reaction times for zero phase difference to be consistently shorter than the times required for *accurate* judgments of near-threshold phase differences. This is probably to be expected on the basis of psychophysical theory if the judgments at zero phase difference are regarded as being toward the negative | 'easily discriminated' end of the function. Fig. 2 shows the reaction time as a function of phase difference in seconds. Functions are traced only through the range of phase differences for which the judgments 'out of phase' were made at least 60 percent of the time. Note that there is a systematic tendency for reaction time to decrease as the phase difference is increased. And, in general, judgments for the faster speeds of oscillation take less time.

TABLE I

MEAN REACTION TIMES IN SECONDS, AND PERCENTAGES JUDGED 'OUT OF PHASE' AT VARIOUS PHASE DIFFERENCES FOR 5 SPEEDS OF OSCILLATION
Means based on 25 readings, 5 readings for each of 5 observers

Phase Difference in Degrees	Speeds of Oscillation									
	52 rpm		72 rpm		288 rpm		710 rpm		830 rpm	
	R.T. in sec.	% out of phase	R.T. in sec.	% out of phase	R.T. in sec.	% out of phase	R.T. in sec.	% out of phase	R.T. in sec.	% out of phase
0	5.6	4	4.8	0	4.5	24	6.6	28	7.0	16
3	8.8	52								
5	6.7	72	6.0	60						
7	6.5	84	6.5	80						
10	5.4	96	4.4	88	5.0	20	6.3	20		
12	4.9	96								
15	3.0	100	2.7	100	6.2	28				
20	3.1	100	2.8	96	5.3	60	6.3	36	8.0	24
25	2.6	100			5.3	64				
30	1.9	100	2.0	100	3.0	92	7.7	36		
35	1.9	100								
40	1.6	100	2.0	100	2.6	100	6.4	40	8.3	48
45							5.8	60		
50	1.7	100	1.5	100	2.2	100	6.2	68		
55							5.6	60		
60	1.6	100	1.4	100	1.6	100	3.9	92	7.5	60
70	1.4	100	1.3	100	1.5	100	3.2	92		
80									5.8	92
90									3.7	96
100	1.3	100	1.4	100	1.3	100	2.7	100	5.7	96
110									3.0	96
120									2.9	100
140							1.3	100	2.9	100
160									2.5	100
180	1.2	100	1.5	100	1.3	100	1.4	100	2.9	100

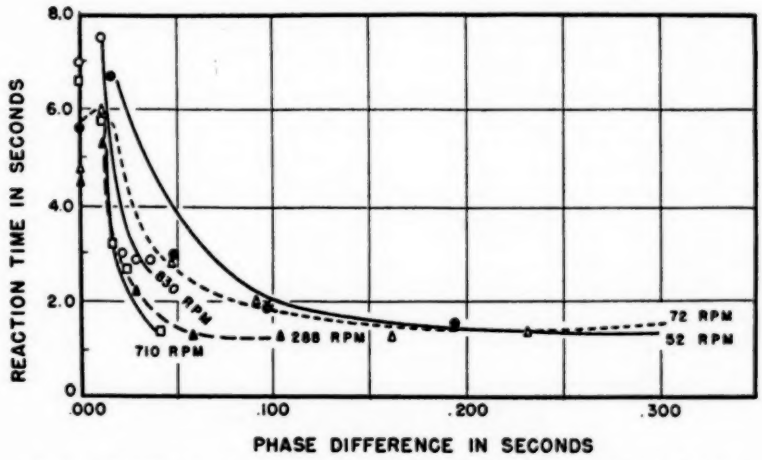

FIG. 2. Reaction time to make a phase discrimination for different amounts of phase-difference in milliseconds. The different curves are for different rates of oscillation.

Fig. 3 is a plot of the percentage of judgments 'out of phase.' This plot includes all data, irrespective of the reaction time. Comparison of Fig. 3 with Fig. 2 shows that the asymptotes for the two kinds of plot, for any frequency, are not reached at the same phase difference. Thus, reaction time continues to decrease for increase in phase difference beyond the earliest point at which the difference can be detected 100 percent of the time.

Fig. 4 is another plot of the reaction-time data in Table I. For the figure, the reaction-time has been converted into the number of oscillations that occurred during the reaction-time period. Thus, the ordinate becomes the reaction-time in number of cycles of oscillation. Whereas Fig. 2 shows the reaction time to be reasonably similar for all speeds, with the faster speeds producing somewhat faster reactions, Fig. 4 shows that the number of stimulations required increases markedly for the faster speeds.

In general, the faster the rate of stimulation, the more quickly can a subject discriminate whether the two stimuli are in or out of phase; and the faster the movement, the less the time equivalent of the minimum perceptible phase difference. This general rule appears to hold up to a certain limiting speed. Thereafter discrimination falls off with further increase in speed. It appears, therefore, that there is an optimum rate of stimulation for quick reliable judgments. Inspection of Fig. 2 for reaction time, and of Table I for the percentage judged 'out of phase,' indicates this optimum to lie *somewhere* between the 288 and 830 rpm conditions.

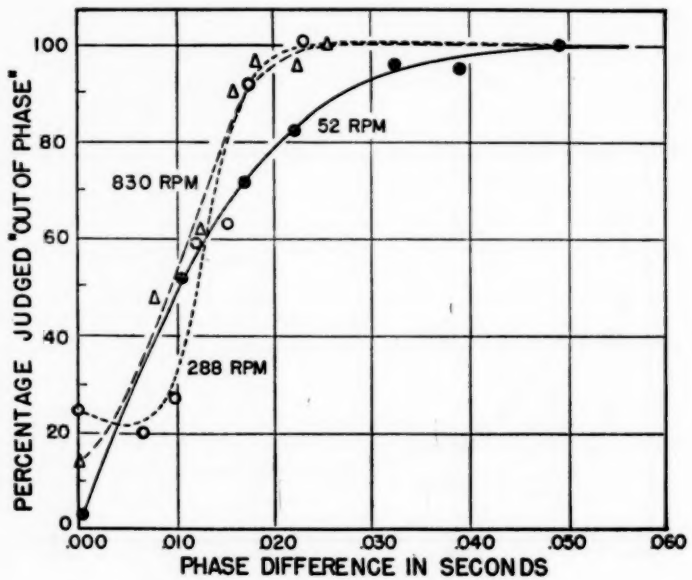

FIG. 3. Percent of differences judged correctly as a function of the phase-difference in seconds

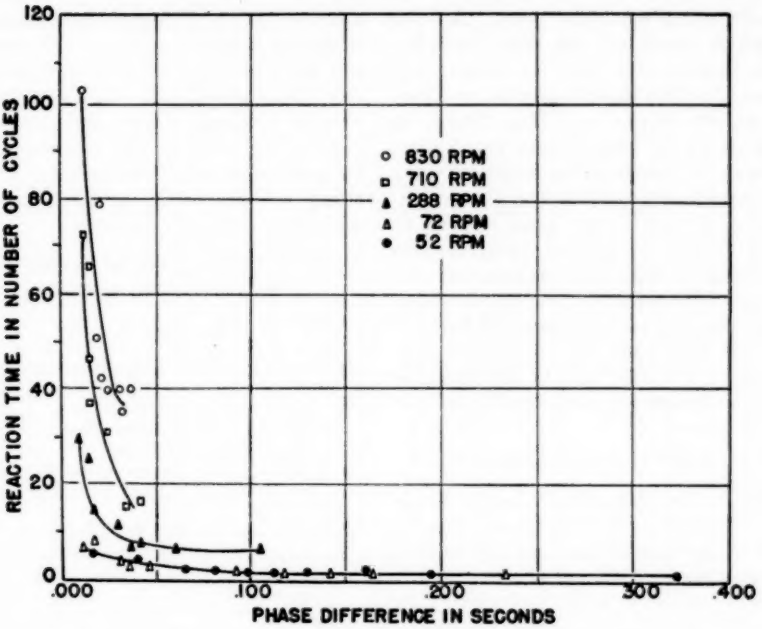

FIG. 4. Reaction time to make a phase discrimination converted into number of cycles of vibration plotted against phase-difference in seconds

DISCUSSION

Before summarizing the conclusions that are indicated by the experiment, several points concerning the results will be touched upon.

It should be pointed out that our technique for assessing reaction time was crude, and therefore the final asymptotic level of the order of 1–1½ sec. in Fig. 2 may be invalid. The same point applies to the end-points for the functions in Fig. 4. Reaction times of the order of 0.5 sec. were the smallest that could be measured with our technique. As a consequence, the higher speeds, even at a phase difference of 180 degrees, appear to have long latencies as measured in number of cycles. In other words, our technique makes these end-points appear to be a direct function of the speed at which the dowels are moving. A more careful analysis of the minimum reaction time, especially for the fast speeds of oscillation, is therefore indicated before the values for the large phase differences can be accepted.

It is of note that in Fig. 2 the function for reaction time against phase difference for the speed of 830 rpm is between those for 72 and 288 rpm. This fact suggests the possibility that a somewhat different sensitivity is involved in the judgments at this highest speed. It was observed that all observers changed their method of 'following' the dowels with their fingertips at the high speed. With the slower speeds, the fingertips, placed on the top surface of the dowels, followed the dowels through

their complete excursions. At the highest speed of 830 rpm, however, the fingers were held in a fixed position, and the dowels merely hit them once per excursion. This change in the mode of stimulating the fingers might have something to do with the receptors involved in the discrimination. For the lower speeds, the whole fingers moved making possible kinesthetic stimulation in muscles and joints. At the high rates of vibration, the stimulation was confined primarily to the pads of the finger tips. How this factor affects the discriminative functions, we do not know.

Summary and Conclusions

Five *O*s were asked to judge whether two vibrating dowels under the finger tips of the first and middle fingers were in phase or out of phase. Their reaction time and their accuracy of judgment were then measured. Speeds up to 830 rpm were used. Phase differences between zero and 180 degrees were studied.

The following are the chief results of this exploratory study: (a) *O*s can perfectly discriminate phase differences of 180 degrees at speeds up to the 830 rpm, the highest used in the study. (b) If three sec. are allowed for judgment, phase differences of the order of 40–50 msec. can be detected accurately at 52 rpm; of 30–40 msec. at 72 rpm; and 17–25 msec. at 288, 710 and 830 rpm. (c) As the time allowed for the judgment to be made decreases, larger phase differences are required for accurate judgment. (d) There appears to be an optimum rate of movement for quick and accurate judgments of phase difference. It was not carefully measured in this study but lies in the general region, probably, of 300–700 rpm.

(Manuscript received for immediate publication April 1, 1949)

References

1. Bellows, R. M. An experimental isolation of some factors determining response to rhythmic cutaneous stimulation. I. Frequency, pressure, and time. *J. exp. Psychol.*, 1936, **19**, 716–731.
2. Geldard, F. A. The perception of mechanical vibration. *J. gen. Psychol.*, 1940, **22**, 243–308.
3. Gilmer, B. von H. The measurement of sensitivity of the skin to mechanical vibration. *J. gen. Psychol.*, 1935, **13**, 42–61.
4. Knudsen, V. O. "Hearing" with the sense of touch. *J. gen. Psychol.*, 1928, **1**, 320–352.
5. Metcalf, J. T. Cutaneous and kinesthetic sense. *Psychol. Bull.*, 1928, **25**, 569–581.
6. Trimble, O. C. The theory of sound localization: a restatement. *Psychol. Rev.*, 1928, **35**, 515–523.
7. Trimble, O. C. Intensity difference and phase-difference as conditions of stimulation in binaural sound localization. *Amer. J. Psychol.*, 1935, **47**, 264–274.
8. Wightman, E. R., & Firestone, F. A. Binaural localization of pure tones. *J. acoust. Soc. Amer.*, 1930, **2**, 271–280.
9. Zigler, M. J. Touch and kinesthesis. *Psychol. Bull.*, 1930, **27**, 298–317.
10. Zigler, M. J. Touch and kinesthesis. *Psychol. Bull.*, 1932, **29**, 260–278.

'SPREAD OF EFFECT' WITHOUT REWARD OR LEARNING

BY FRED D. SHEFFIELD

Yale University

PURPOSE

In a previous study (5) evidence was obtained that Thorndike's (8, 9) spread of effect is an artifact not directly related to the reinforcing effect of reward. The phenomenon was found instead to be understandable in terms of guessing habits of the subjects, the role of reward apparently being only to insure more repetition of rewarded than of unrewarded responses. The present experiment was undertaken to provide a crucial test of the interpretation that a 'spread of effect' is produced by guessing habits alone. It is assumed that the reader will be fairly familiar with Thorndike's 'spread of effect' and its significance for Thorndike's theoretical position, so no summary of the literature is included here. For a recent review, see Hilgard (2).

METHOD

The experiment avoided all reward or learning; it tested for Thorndike's phenomenon with *chance* repetitions in the course of guessing sequences. According to the guessing-sequence explanation one should find the phenomenon whether repetition is caused by learning of correct responses or whether repetition is simply due to chance.

The experiment was conducted with a group of 155 Yale undergraduates in a second-term course in general psychology. The students were told that the purpose was to check on the 'clairvoyance' findings at Duke University. They were asked to try to divine by clairvoyance the order of digits in four specified columns of Tippett's (11) tables of random numbers. It was hoped that the use of this ruse would make the guessing of numbers seem sensible to the students. Guesses were recorded by the subjects on mimeographed sheets having four columns of 50 spaces each. The following instructions were read:

> You are going to try and guess the numbers in four of the columns from this table of random numbers. We will do one column at a time and will not start a new column until all have completed the preceding one. I have selected, as the first column that you will guess, the last column of the first page in the table of random numbers. You will record your guesses for this column in the 50 spaces provided under the letter 'A' on your answer sheet. Make your guesses in order, that is, guess the first number in the column, then the second, and so on, down to the 50th and last number. If there is no clairvoyance your guesses will be right only by chance—to the extent that clairvoyance operates you will guess the numbers printed in the table more often than chance allows. It may help to try and visualize the numbers as if seeing them printed in the table. In any case always write down your first impression and move along fairly rapidly. You should spend no more than 2 or 3 seconds per number. That is, it should take only about two minutes to fill in column A. When you have finished A, wait, do not go on to column B. Remember you are to guess only single-digit numbers ranging from 0 through 9. These are in a random order in the table and you are to try and divine this order. Are there any questions?

Reading these instructions led to unforeseen questions about the exact meaning of 'random numbers,' and the experimenter was obliged to extemporize a few additional sentences in which it was stated that a random sequence of the 10 digits is one in which chance alone operates and each digit is equally likely to appear at any point. In addition the students were asked to visualize successive drawings from a large container filled with a well-stirred mass of small and uniform capsules, each of which contained one of the 10 digits, there being the same number of capsules for each digit.

The instructions for the remaining three columns, given as each column was successively completed, were as follows:

Column B

> Are there any questions before we go on to the next column? Before we go on to the next column take a sheet of paper and cover up your guesses in column A. For column B,

I have selected the last column of the third page in the table of random numbers. You may begin column B now.

Column C

Is everyone finished? Then start in on C, for which I have selected the last column of the fifth page of random numbers. Keep both A and B covered up with your sheet of paper.

Column D

Is everyone finished? Then do column D, which is the last column of the seventh page in the table. Keep A, B, and C well covered.

Results

Analysis of results was made as similar as possible to that used in the earlier spread of effect studies. The four columns of 50 guesses each were treated as if they were responses on four presentations of a list of 50 stimuli. Each repetition of the same guess in the same ordinal position in two successive columns was scored as a 'repeat' and the presence or absence of other repeats in the five *following* positions was determined. Analysis was restricted to following positions because the present procedure necessarily produces a symmetrical 'before' and 'after' gradient. If, for example, two repeats are separated by one non-repeat, the first repeat is two 'before' and the second is two 'after.' This differs from Thorndike's studies, in which rewarded and unrewarded repeats are distinguished. In the present study repeats are undifferentiated: all are unrewarded so there are no reward points to separate before and after gradients. The present analysis is essentially a test of whether repeats tend to come in *clusters* as implied by the guessing-sequence hypothesis.

To facilitate analysis the guesses were punched on IBM cards, there being four cards of 50 guesses each per subject. By superimposing pairs of cards, the points at which repeats occurred readily appeared as joint punches.

Findings are shown in Table I, which gives the percent of repeats in the designated positions after each chance repeat. It can be seen that a descending gradient of mean percentage of repeats was obtained, in which the high point is in the position *one* after, and the low point—approximately chance—is obtained in the fifth position following a repeat.

TABLE I

Gradient of Repetition Following Chance Repeats

	Position After Chance Repeat				
	1	2	3	4	5
Percent Repetition	13.3	11.4	11.2	10.6	10.2

A statistical problem in most spread of effect studies is that the N on which percent of repetition is based is the number of *responses*, whereas the experimenter wishes to generalize to his population of *subjects*. This could be avoided by giving each subject a percentage-of-repetition score at each position before or after reward and averaging these scores over all subjects. The difficulty with this procedure is that such scores tend to have a very skewed distribution which makes the application of normal statistics questionable. A given subject may show only a few rewarded repeats, and the modal unrewarded repeat score tends to be close to zero. To circumvent this difficulty in the present analysis, the 155 subjects were broken down into 15 random sets of 10 each, five being discarded at random to bring the N down to 150. Percentage of repetition was then computed in each sub-group in the usual way—in which N is the number

of responses. This gave fairly symmetrically distributed sample percentages and 14 degrees of freedom to evaluate mean differences between the different positions following a repeat. In effect the population sampled consisted of successive independent *experiments* of 10 subjects each.

Using this procedure the difference in percent repetition between positions one and two after reward had a *P*-value of .03, that between position one and all other points had a *P*-value of less than .01.

Discussion of Results

The chief conclusion is that the guessing-sequence hypothesis is confirmed under circumstances in which differential effects of reward were absent. The results give positive proof of the operation of the guessing-sequence factor in producing an 'after' gradient. It is believed that the results demonstrate a factor sufficient to explain entirely Thorndike's gradient, in terms that do not assume the automatic reinforcing effect of reward postulated by Thorndike (8, 9).

A discrepancy between the present results and what might be expected from the guessing-sequence hypothesis is the relatively low level of repetition despite the fact that repetition in the zero position was 100 percent. It seems probable that one factor contributing to this is the lack of any stimulus terms, so that word-number associations and incidental learning of chance guesses do not raise the level of repetition. Another possible factor is that the subjects were self-consciously trying to make a random distribution of numbers. The nature of random numbers was probably over-emphasized in the instructions, and, since the subject could keep track of his guesses within any column of 50, he might be less subject to influence from previous guesses. Further, the subjects were not distracted by any learning task and may therefore have been less prone to guess in automatic sequences.

Despite the questions raised by the low level of repetition, the guessing-sequence factor was demonstrated to be sufficient to produce an 'after' gradient. It would not, of course, produce the 'before' gradient generally considered part of Thorndike's phenomenon. However, the evidence for a true 'before' gradient is completely lacking. In two of the most extensive studies of the phenomenon —by Brandt (1) and by Zirkle (12, 13) —the 'before' gradients have about zero slopes. Before gradients demonstrated in comparisons using only rewarded *repeats* (5, 7) are subject to an artifact of selection previously described (5). Further, Tilton (10) demonstrated an artifact by which the procedures used by Thorndike could cause the before gradient to be an artifact of the after gradient. And Hilgard (2) has recently pointed out a previously unnoticed artifact which contributes to both gradients, at least in some of the earlier studies in which many repetitions of lists were used. It seems fair to say that an 'after' gradient is well established but that there is no proof of a 'before' gradient. It is this 'after' gradient which can be explained by the subjects' guessing sequences.

Zirkle (12, 13) has presented findings which he interprets as evidence that Thorndike's phenomenon is a spread of effect of the isolation factor described by von Restorff (6). According to Zirkle, saying 'right' isolates the rewarded response, making it better remembered. He believed the isolation factor spreads to neigh-

boring responses because of their 'association' with the rewarded response. However, Zirkle's theory does not explain a number of the relevant facts, including his own finding (12), that if stimuli are rearranged in successive trials the correct response is attached to its stimulus but the incorrect adjacent responses are not attached to theirs. Instead the spread is found after the correct *response* in its new position. If isolation accounts for the connecting of the correct response to its stimulus it cannot account for the spread, in which responses *fail* to be connected to their stimuli. Furthermore, findings subsequent to Zirkle's presentation of his theory are negative. Isolation does not account for the fact that spread is found only for rewarded *repeats* (3, 5, 7), and in a direct test for the spread of isolation in serial learning Jenkins and Postman (4) found the von Restorff effect on isolated items but a negative effect on adjacent items.

On the other hand, Zirkle's results with isolation can be readily explained by the guessing sequence factor, assuming only the well known fact that isolation assists learning of the correct response and thereby makes it more frequently repeated. Zirkle found that increasing isolation increased spread, but isolation also always increased repetition of the key response. Furthermore Zirkle's experiment with rearranged stimuli, which is not explained by his theory, is strong positive evidence for the guessing-sequence explanation. This result has recently been confirmed, with adequate statistical tests, by Jenkins and Cunningham (3).

Most, if not all, of the major findings in studies of the spread of effect are readily explained in terms of the guessing-sequence factor. This was shown for earlier findings in a previous study (5). The more recent important findings of Zirkle (12, 13) have been shown in the present discussion to be in line with or in support of the guessing-sequence interpretation. The same is not true of any other interpretation. This fact is not made clear in Hilgard's review (2) of studies in this field, which treats guessing-sequences as a possible factor rather than as the *one* most parsimonious explanation of all of the relevant facts. It is believed that the present study provides some new relevant facts that can be explained in no other way.

Summary

1. Subjects were asked to try to divine—through clairvoyance—the sequence of random numbers in four specified columns of Tippett's tables.

2. The guessing sequences obtained were subjected to a spread-of-effect analysis in which chance repeats in the same ordinal position in successive columns were treated as if they were rewarded responses in a spread study.

3. The result was a reliable 'after' gradient such as is found in spread-of-effect studies. The findings confirm the 'guessing-sequence' explanation of Thorndike's phenomenon and provide a 'spread of effect' that cannot be explained by the alternative theories of Thorndike and others.

(Manuscript received for immediate publication May 26, 1949)

References

1. Brandt, H. The spread of the influence of reward to bonds remote in sequence and time. *Arch. Psychol., N. Y.*, 1935, **27**, No. 180.
2. Hilgard, E. R. *Theories of learning.* New York: Appleton-Century, 1948.
3. Jenkins, W. O., & Cunningham, Leta M. The guessing-sequence hypothesis, the 'spread of effect,' and number-guessing habits. *J. exp. Psychol.*, 1949, **39**, 158–168.

4. Jenkins, W. O., & Postman, L. Isolation and 'spread of effect' in serial learning. *Amer. J. Psychol.*, 1948, 61, 214–221.
5. Jenkins, W. O., & Sheffield, F. D. Rehearsal and guessing habits as sources of the 'spread of effect.' *J. exp. Psychol.*, 1946, 36, 316–330.
6. Restorff, Hedwig von. Analyse von Vorgängen im Spurenfeld. I. Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychol. Forsch.*, 1933, 18, 299–342.
7. Taylor, D. W. Spread of effect in one-trial learning $\frac{7}{8}$ in relation to $\frac{1}{8}$ the guessing-sequence $\frac{1}{3}$ hypothesis. *Amer. Psychologist*, 1947, 2, 313.
8. Thorndike, E. L. A proof of the law of effect. *Science*, 1933, 77, 173–175.
9. Thorndike, E. L. An experimental study of rewards. *Teach. Coll. Contr. Educ.*, 1935, No. 580.
10. Tilton, J. W. Gradients of effect. *J. genet. Psychol.*, 1945, 66, 3–19.
11. Tippett, L. H. C. *Random sampling numbers.* London: Cambridge Univ. Press, 1927.
12. Zirkle, G. A. Success and failure in serial learning: I. The Thorndike effect. *J. exp. Psychol.*, 1946, 36, 230–236.
13. Zirkle, G. A. Success and failure in serial learning: II. Isolation and the Thorndike effect. *J. exp. Psychol.*, 1946, 36, 302–315.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The issues listed below are available for sale

Year	Volume	Available Numbers						Price per Number	Price per Volume
1916	1	1	2	3	4	5	6	$1.25	$7.00
1917	2	1	2	–	4	5	–	$1.25	$5.00
1920	3	–	2	3	4	5	6	$1.25	$6.25
1921	4	1	2	3	4	5	6	$1.25	$7.00
1922	5	1	–	3	4	5	6	$1.25	$6.25
1923	6	–	–	–	–	–	–	——	——
1924	7	1	2	–	–	5	–	$1.25	$3.75
1925	8	1	2	3	4	–	6	$1.25	$6.25
1926	9	–	2	3	–	5	6	$1.25	$5.00
1927	10	–	–	3	4	5	6	$1.25	$5.00
1928	11	1	2	3	4	5	6	$1.25	$7.00
1929	12	1	2	3	4	5	6	$1.25	$7.00
1930	13	1	2	3	4	5	6	$1.25	$7.00
1931	14	1	2	3	4	5	6	$1.25	$7.00
1932	15	1	2	3	4	5	6	$1.25	$7.00
1933	16	1	2	3	4	5	6	$1.25	$7.00
1934	17	1	2	3	4	5	6	$1.25	$7.00
1935	18	1	2	3	4	5	6	$1.25	$7.00
1936	19	1	2	3	4	5	6	$1.25	$7.00
1937	20	1	2	3	4	5	6	$1.25	$7.00
	21	1	2	3	4	5	6	$1.25	$7.00
1938	22	1	2	3	4	5	6	$1.25	$7.00
	23	1	2	3	4	5	6	$1.25	$7.00
1939	24	1	2	3	4	5	6	$1.25	$7.00
	25	1	2	3	4	5	6	$1.25	$7.00
1940	26	1	2	3	4	5	6	$1.25	$7.00
	27	1	2	3	4	5	6	$1.25	$7.00
1941	28	1	2	3	4	5	6	$1.25	$7.00
	29	1	2	3	4	5	6	$1.25	$7.00
1942	30	1	2	3	4	5	6	$1.25	$7.00
	31	1	2	3	4	5	6	$1.25	$7.00
1943	32	1	2	3	4	5	6	$1.25	$7.00
	33	1	2	3	4	5	6	$1.25	$7.00
1944	34	1	2	3	4	5	6	$1.25	$7.00
1945	35	1	2	3	4	5	6	$1.25	$7.00
1946	36	1	2	3	4	5	6	$1.25	$7.00
1947	37	1	2	3	4	5	6	$1.25	$7.00
1948	38	1	2	3	4	5	6	$1.25	$7.00
1949	39	By subscription, $7.00						$1.25	——
		List price, Volumes 1 through 38							$247.50
		30% discount							74.25
		Net price, Volumes 1 through 38							$173.25

Information about the Journal of Experimental Psychology: as noted in the table, the journal was not published in 1918 and 1919; from 1937 to 1943 it was published at the rate of two volumes per year. Eighteen numbers affecting eight volumes are out of print in the entire series.

Information about prices: when all numbers are available, the price per volume throughout the series is $7.00. For incomplete volumes, the price is $1.25 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Address orders and subscriptions to

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue, N.W. Washington 5, D. C.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue, N.W.
Washington 5, D. C.

The American Psychological Association, founded in 1892 and incorporated in 1925, is the major psychological organization in the United States. With approximately 6,700 members and 2,000 affiliates, it includes most of the qualified psychologists in the country. The purpose of the APA is to advance psychology as a science, as a profession, and as a means of promoting human welfare. It attempts to further these objectives by holding annual meetings, publishing psychological journals, conducting a personnel placement service for psychologists, and working toward improved standards for psychological training and service.

In order to give recognition to the specialized interests of different psychologists, the APA includes seventeen Divisions. Any person, after becoming a member of the APA, may apply for membership in as many Divisions as he wishes. The Divisions are:

Division of General Psychology
Division on the Teaching of Psychology
Division of Experimental Psychology
Division on Evaluation and Measurement
Division on Childhood and Adolescence
Division of Personality and Social Psychology
The Society for the Psychological Study of Social Issues—a Division of the American Psychological Association
Division on Esthetics
Division of Clinical and Abnormal Psychology
Division of Consulting Psychology
Division of Industrial and Business Psychology
Division of Educational Psychology
Division of School Psychologists
Division of Counseling and Guidance Psychologists
Division of Psychologists in Public Service
Division of Military Psychology
Division on Maturity and Old Age

Each Division has its own officers. Each meets annually at the time and place of the APA meeting. Each has its own membership requirements, which in some cases are higher or more specialized than the requirements for election to the APA.

The annual meetings of the APA are held in September. An attempt is made to choose meeting sites in such a way that attendance will be convenient for members and their guests in different sections of the country at different times. Thus, the 1949 meeting will be held on September 6 to 10 in Denver, and the 1950 meeting on September 4 to 11 at Pennsylvania State College. These meetings provide the members with an opportunity to present or hear reports of psychological research and lectures on topics of psychological interest, and to participate in formal and informal discussion groups. The research reports, lectures, and discussion groups cover varied topics of current interest within the field of psychology.

The chief governing body of the APA is the Council of Representatives. It includes representatives from each of the Divisions. A Board of Directors, composed of the four officers of the APA (President, President-elect, Recording Secretary, and Treasurer) and six Council members elected by the Council, is the administrative agent of the Council and exercises general supervision over the affairs of the Association.

The APA maintains an office in Washington which acts as a coordinating center for all APA activities. Journal subscriptions, membership applications, personnel placement requests, and general APA business procedures are handled in this office.